

Г. А. Кухарев

# БИОМЕТРИЧЕСКИЕ СИСТЕМЫ

МЕТОДЫ И СРЕДСТВА ИДЕНТИФИКАЦИИ  
ЛИЧНОСТИ ЧЕЛОВЕКА







**TECHNICAL UNIVERSITY  
OF SZCZECIN  
FACULTY OF COMPUTER  
SCIENCE  
AND INFORMATION SYSTEMS**



Żołnierska Street 49, 71-210  
Szczecin,

Poland, tel./fax: (+48-91) 449-55-40, e-mail: wi@wi.ps.pl,  
<http://www.wi.ps.pl>

Dean of Computer Science and Information Systems Faculty  
**prof. Jerzy Soldek**

---

***Employment:***

- 19 professors, • 27 doctors of science, • 61 masters of science,
- 36 persons of technical, administration and services staff
- The Faculty confers the degree of PhD in Computer Science

**Research teams:**

- |   |   |
|---|---|
| • Computer Graphics and Multimedia        | • Supercomputing and Parallel Programming   |
| • Computer Aided Decisions                | • Computer and Telecommunication Networks   |
| • Banking Information Technology          | • Cryptography and Software Engineering     |
| • Artificial Intelligence, Expert Systems | • Data Bases and Information Systems Design |
| • Image Recognition and Processing        | • Information Systems                       |
| • Multivalued Logic                       |   |
| • Synthesis of Reprogrammable Systems     |   |
| • Mathematical Modelling                  |   |

**Conferences organized by faculty:**

- Advanced of Computer Systems – ACS
- Pattern Recognition and Information Processing — PRIP
- Reconfigurability Unification & Customization — RUC





**TECHNICAL UNIVERSITY  
OF SZCZECIN  
FACULTY OF COMPUTER  
SCIENCE AND INFORMATION  
SYSTEMS**



Żołnierska Street 49,  
71-210 Szczecin,  
Poland, tel./fax: (+48-91) 449-55-40, e-mail: [wi@wi.ps.pl](mailto:wi@wi.ps.pl),  
<http://www.wi.ps.pl>

Dean of Computer Science and Information Systems Faculty  
**prof. Jerzy Soldek**

---

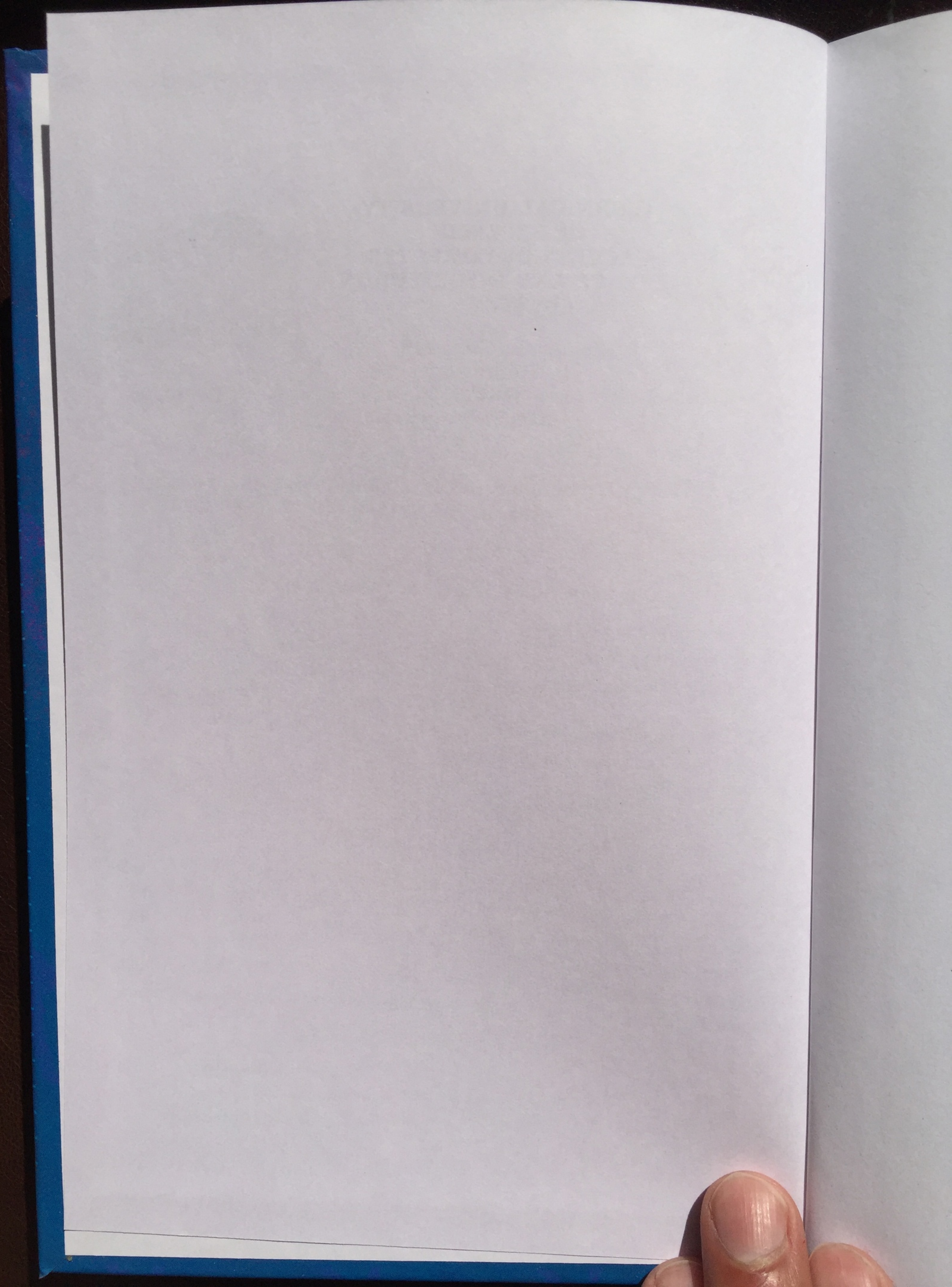
**Diploma Courses (Pedagogic Specializations):**

- Software Engineering
- Microprocessor Systems
- Computer and Telecommunication Networks
- Computer Graphics and Multimedia
- Teledetection and Geographical Information Systems
- Information Systems
- Real Time Systems
- Systems of Computer Aided Decision
- Computer Systems in Administration
- Print and Publishing Systems

**Cooperation:**

- **France:** University of Technology (Troyes), Ecole Centrale de Lyon (Lyon) and Project **STAR** – Specialized Training in Aeronautic Research (Merignac, Bordeaux)
- **Canada:** Saint Francis Xavier University
- **Germany:** University of Goetting
- **Russian:** Electrotechnical University of St. Petersburg
- **Belarus:** Academy of Science and State University of Informatics and Radioelectronics







Посвящается  
моему первому учителю  
профессору  
**Сергею Александровичу Майорову**



Посвящение  
моему любимому сыну  
профессору  
Сергию Александровичу Мейерову

БИ



Г. А. КУХАРЕВ

# БИОМЕТРИЧЕСКИЕ СИСТЕМЫ

## МЕТОДЫ И СРЕДСТВА ИДЕНТИФИКАЦИИ ЛИЧНОСТИ ЧЕЛОВЕКА



**ПОЛИТЕХНИКА  
ИЗДАТЕЛЬСТВО**  
Санкт-Петербург 2001



УДК 612-087  
ББК 67.99(2)8  
К96

Федеральная программа книгоиздания России

**Кухарев Г. А.**

**К96** Биометрические системы: Методы и средства идентификации личности человека.— СПб.: Политехника, 2001.— 240 с., ил.

ISBN 5-7325-0623-3

Дан обзор идей, методов и средств, используемых в системах идентификации и верификации личности человека на основе его биометрических характеристик. Проанализированы состояние и тенденции развития различных биометрических технологий. Подробно обсуждены биометрические системы, использующие информацию об отпечатках пальцев, геометрии руки, параметрах лиц и радужной оболочке глаз. Представлены элементы цифровой обработки изображений, наиболее часто используемые в различных биометрических технологиях, базовые методы обработки исходных данных в приложении к системам распознавания лиц, а также методики построения (синтез, анализ и реализация) систем распознавания людей по их лицам. Приведены словарь терминов, обширная библиография и адреса наиболее важных интернет-сайтов по биометрическим технологиям.

Предназначена для студентов и аспирантов соответствующих специальностей, а также практиков, интересующихся использованием представленных биометрических систем.

УДК 612-087  
ББК 67.99(2)8

ISBN 5-7325-0623-3



НАУЧНОЕ ИЗДАНИЕ

**Кухарев Георгий Александрович**  
**БИОМЕТРИЧЕСКИЕ СИСТЕМЫ**  
**МЕТОДЫ И СРЕДСТВА**  
**ИДЕНТИФИКАЦИИ ЛИЧНОСТИ ЧЕЛОВЕКА**

Редактор *Т. И. Ильичева*. Переплет *Н. В. Вавулина*. Технический редактор *Т. М. Жилич*. Корректоры *Т. Н. Гринчук*, *З. С. Романова*. Верстка *А. В. Гаврилов*.

ЛР № 010292 от 18.08.98 г.

Сдано в набор 22.06.01. Подписано в печать 22.11.01. Формат издания 60×90<sup>1</sup>/<sub>16</sub>. Бумага офсетная. Гарнитура SchoolBook. Печать офсетная. Усл. печ. л. 15,0. Уч.-изд. л. 15,9. Тираж 1000 экз. Заказ № 3659.

Государственное предприятие «Издательство „Политехника“».  
191011, Санкт-Петербург, ул. Инженерная, д. 6.

Отпечатано с готовых диапозитивов в ГП «Типография им. П. Ф. Анохина».  
185005, г. Петрозаводск, ул. «Правды», д. 4.

ISBN 5-7325-0623-3

© Г. А. Кухарев, 2001



# О Г Л А В Л Е Н И Е

|   |            |
|---|------------|
| Предисловие .....   | 5          |
| <b>1. Введение в биометрику .....</b>   | <b>7</b>   |
| 1.1. Вводные понятия .....  | 8          |
| 1.2. Биометрические характеристики и их свойства .....                          | 13         |
| <b>2. Обзор некоторых биометрических технологий .....</b>                       | <b>18</b>  |
| 2.1. Технология отпечатков пальцев. История применения .....                    | 18         |
| 2.1.1. Информационные признаки отпечатков пальцев .....                         | 19         |
| 2.1.2. Стандарты на отпечатки пальцев .....                                     | 20         |
| 2.1.3. Устройства для получения отпечатков пальцев .....                        | 23         |
| 2.1.4. Сравнение отпечатков пальцев .....                                       | 27         |
| 2.1.5. Несколько слов о других подходах к реализации AFIS .....                 | 33         |
| 2.2. Технология, использующая геометрию руки .....                              | 38         |
| 2.3. Технология, использующая распознавание лиц .....                           | 42         |
| 2.3.1. Обзор методов распознавания лиц .....                                    | 44         |
| 2.3.2. База данных изображений лиц .....  | 46         |
| 2.3.3. Введение в методы PCA, ПКЛ и LDA .....                                   | 47         |
| 2.3.4. Отбор признаков при распознавании лиц .....                              | 52         |
| 2.3.5. Анализ локальных признаков .....   | 55         |
| 2.3.6. Использование эластичных моделей форм лица .....                         | 58         |
| 2.3.7. Коммерческие системы .....   | 61         |
| 2.3.8. Несколько слов о будущем .....   | 65         |
| 2.4. Технология, использующая распознавание по радужной оболочке глаза .....    | 67         |
| <b>3. Элементы цифровой обработки изображений .....</b>                         | <b>73</b>  |
| 3.1. Формы и способы представления цифровых изображений .....                   | 73         |
| 3.2. Система обозначений, используемая в алгоритмах обработки изображений ..... | 77         |
| 3.3. Базовые процедуры обработки изображений .....                              | 78         |
| 3.4. Инварианты геометрических преобразований .....                             | 85         |
| 3.5. Обработка изображений с использованием преобразований Фурье .....          | 93         |
| 3.5.1. Изображение и его преобразование Фурье .....                             | 93         |
| 3.5.2. Применение дискретного преобразования Фурье .....                        | 95         |
| 3.6. Метод главных компонент .....  | 103        |
| 3.7. Преобразование Фурье—Меллина .....   | 107        |
| <b>4. Предварительная обработка исходных данных .....</b>                       | <b>116</b> |
| 4.1. Выделение изображения лица, его деталей и параметров .....                 | 116        |
| 4.1.1. Выделение лица и его частей .....  | 116        |
| 4.1.2. Определение расстояния между центрами глаз .....                         | 122        |
| 4.1.3. Выделение изображения лица градиентными методами .....                   | 124        |
| 4.2. Согласование размеров изображений .....                                    | 125        |
| 4.3. Нормализация образов .....   | 130        |
| 4.3.1. Элиминирование поворотов .....   | 130        |
| 4.3.2. Элиминирование наклонов .....  | 132        |



|  |            |
|--|------------|
| 4.3.3. Нормализация яркости изображений . . . . .                                    | 135        |
| <b>5. Синтез системы распознавания для структурированных данных</b> . . . . .        | <b>141</b> |
| 5.1. Автоматическое разделение данных на обучающую<br>и контрольную группы . . . . . | 141        |
| 5.2. Отбор первичных признаков и их редукция . . . . .                               | 143        |
| 5.3. Уменьшение пространства признаков . . . . .                                     | 146        |
| 5.4. Процесс распознавания контрольных образов . . . . .                             | 149        |
| 5.5. Выбор параметров системы распознавания . . . . .                                | 151        |
| 5.5.1. Изменение числа классов и числа образов в классе . . . . .                    | 151        |
| 5.5.2. Распознавание при наложении шума . . . . .                                    | 158        |
| 5.5.3. Распознавание при масштабировании образов . . . . .                           | 161        |
| 5.5.4. Распознавание при масштабировании и наложении<br>шума . . . . .               | 164        |
| 5.5.5. Распознавание образов, не принадлежащих собствен-<br>ной базе . . . . .       | 169        |
| <b>6. Системы распознавания «on-line»</b> . . . . .                                  | <b>172</b> |
| 6.1. Структура системы распознавания . . . . .                                       | 172        |
| 6.2. Виртуальные базы данных и выбор признаков . . . . .                             | 179        |
| 6.3. Синтез системы «Visitor Identification» . . . . .                               | 194        |
| 6.4. Системы, обеспечивающие контроль доступа . . . . .                              | 208        |
| 6.5. Интегрированные системы распознавания . . . . .                                 | 222        |
| Тестовая база данных (ORL Database of Faces) . . . . .                               | 225        |
| Словарь терминов * . . . . .   | 229        |
| Литература . . . . .   | 234        |



## ПРЕДИСЛОВИЕ

Вниманию читателей предлагается монография, в которой сделана попытка представить основные идеи, раскрыть методы и средства, используемые в настоящее время в биометрических технологиях идентификации личности человека, а также рассмотреть тенденции развития и применения биометрики в будущем.

Одно из перспективных и бурно развивающихся направлений биометрических технологий связано с распознаванием человеческих лиц. В монографии представлены известные на сегодня методы и средства решения этой задачи, которые доведены до практической реализации. Две отдельные главы посвящены практическим работам, связанным с построением, синтезом и анализом систем распознавания лиц, работающих в режимах «off-line» и «on-line».

Монография состоит из шести основных разделов.

**Первый раздел** посвящен рассмотрению биометрических технологий и биометрических систем, их характеристик, а также тенденций на рынке продаж. Здесь же приведены адреса сайтов, на которых можно найти дополнительную информацию о коммерческих биометрических продуктах и системах, а также фирмах — производителях биометрических технологий.

**Второй раздел** посвящен обзору методов и средств биометрических технологий, использующих отпечатки пальцев, геометрию руки, очертаний лиц и особенностей радужной оболочки глаза для идентификации личности. Здесь же дана оценка состояния и перспектив развития соответствующих направлений биометрических технологий, в некоторых подразделах представлены и коммерческие системы — лидеры в своих направлениях.

**Третий раздел** посвящен цифровой обработке изображений и тем методам, которые наиболее часто используются в биометрических технологиях обработки исходной информации. Здесь охарактеризованы формы и способы отображения цифровых изображений, базовые процедуры обработки и наиболее часто используемые в задачах распознавания образов виды преобразований — Фурье, Фурье—Меллина и Карунена—Лоэва. Для каждого из представленных методов обработки приведена соответствующая программа в языке пакета MATLAB.

В **четвертом разделе** представлены методы предварительной обработки исходных данных для распознавания лиц. Раздел включает практические решения задачи автоматического выделения образа лица из исходных изображений (сцен), деталей лица, а также определение антропометрических параметров лиц. В разделе приведены методы и конкретные алгоритмы согласования размеров изображений лиц, а также процедуры нормализации изображений лиц (элиминирование наклонов и поворотов головы, нормализация яркости исходных изображений).



**Пятый раздел** посвящен методике построения (синтез и анализ) систем распознавания лиц на основе использования баз данных с лицами. Здесь рассмотрены: решение задачи разделения исходных данных на обучающую и контрольную группы, методика отбора первичных признаков и способы их редукции, собственно процесс распознавания (сравнения с эталоном) и методика выбора параметров системы распознавания. Раздел иллюстрирован практическими решениями и примерами.

**В шестом разделе** проанализированы системы «on-line» распознавания лиц, которые работают в режиме реального времени с видеокамерой. Обсуждаются возможные варианты построения подобных систем распознавания, их структура и способы повышения точности распознавания. Подробно обсуждаются два класса систем «on-line»: предназначенные для идентификации посетителей и контроля доступа. Раздел иллюстрирован примерами из практики решения задач распознавания лиц в режиме «on-line».

В заключение приведены изображения из тестовой базы лиц и словарь терминов по биометрике.

Монография появилась в ходе подготовки курса лекций и лабораторных работ по биометрическим технологиям на факультете информатики Щецинского политехнического университета (Польша). Экспериментальная часть представленного в ней материала выполнена в лаборатории «Обработки изображений и распознавания образов», входящей в состав факультета информатики, а также при реализации коммерческого проекта для американской фирмы ALDEC.

Автор выражает искреннюю благодарность декану факультета информатики проф. Е. Солдеку (J. Soldek) за организацию лаборатории «Обработки изображений и распознавания образов» и постоянную заботу о ней, что позволило подготовить настоящую монографию и на практике решать задачи биометрики.

Автор считает приятным долгом поблагодарить также президента фирмы ALDEC С. Хайдука (S. Hyduke) за инициализацию интереса к биометрике и финансовую поддержку научных исследований по распознаванию лиц.

Наконец, автор считает своим приятным долгом поблагодарить сотрудников лаборатории «Обработки изображений и распознавания образов» П. Форчманского (P. Forczmański), М. Боравского (M. Borawski), Е. Полролничка (E. Pórolniczak), Т. Пониковского (T. Ponikowski), а также А. Туяку (A. Tujaka), Т. Брылку (T. Bryłka) и многих своих студентов, чьи исследовательские работы и научные дискуссии с которыми позволили получить положительные практические результаты, а также способствовали улучшению методики изложения материалов настоящей монографии.

Фотопортреты лиц, которые использованы в монографии, принадлежат в большинстве случаев сотрудникам факультета информатики и лаборатории «Обработки изображений и распознавания образов» и публикуются с их разрешения.

Отзывы о монографии можно присылать по электронному адресу: [georgi.kucharev@wi.ps.pl](mailto:georgi.kucharev@wi.ps.pl)



## 1. ВВЕДЕНИЕ В БИОМЕТРИКУ

В 1974 г. вышла в свет монография «Pattern Recognition Principles» («Принципы распознавания образов»)\*, авторы которой Ту и Гонсалес отметили одну из нарождающихся тенденций тех лет: «...В нашем обществе создается такое положение, при котором все большее количество людей и организаций занимаются обработкой информации и все меньшее — обработкой материальных объектов. Ощутимой стала потребность в более совершенных информационных системах...» [Тou74].

За прошедшие 26 лет эта тенденция интенсивно и экстенсивно развивалась, в результате чего человечество перешло от простых информационных систем к использованию интеллектуальных сред — вычислительных и информационных технологий современного поколения [Пен00].

Примерами интеллектуальных сред являются [Sol97, Пен00]:

- интернет, через который происходит, в частности, общение людей (видеоконференции, консультации, почта), поиск информации и доступ к ней, реализуются интернет-маркетинг и интернет-магазины и т. д.;

- виртуальные предприятия, реализующие проекты по типу «start up», сообщества и ассоциации, осуществляющие обмен информацией, технологиями, другими информационными и вычислительными ресурсами;

- высокоорганизованные интерфейсы систем «клиент—банк»;

- системы класса «Visitor Identification», реализующие распознавание или аутентификацию, например, постоянного покупателя магазина, гостя (посетителя) частного дома или офиса, пассажира международного транспорта и т. д.;

- системы контроля доступа к информации (базам данных), к средствам (техническим, финансовым), к помещениям, к ресурсам (например, вычислительным);

- системы поиска фотографий лиц людей в базах данных по заданному образцу (например, фотороботу).

Для всех перечисленных примеров характерно одно: все задачи решаются в процессе непосредственного общения человека со средой (в режиме «on-line») и, что самое важное, — этим человеком может быть только зарегистрированный пользователь, подлинность личности которого проверяется на входе в интеллектуальную среду. Поэтому, очевидно, что одной из актуальных проблем широкого использования интеллектуальных сред является проблема идентификации (аутентификации) конечного пользователя.

На современном этапе развития интеллектуальных сред задачи идентификации (аутентификации) конечного пользователя решаются в рамках биометрических технологий [Bio\*\*, Mat00, Пан00].

\* В России ее перевод был опубликован в 1978 г. — Прим. авт.



Так, например, в мае 2000 г. компания Microsoft заявила [Mat00] о том, «что все будущие версии ОС Windows будут иметь встроенные средства поддержки биометрического интерфейса (Biometric Application Programming Interface — BAPI) и технологии авторизации SecureSuite. Совместно с компанией I/O Software [Bio\*\*], занимающей лидирующие позиции в области программных средств биометрики, Microsoft намеревается реализовать комплексный подход к идентификации пользователей во всех своих последующих операционных системах...»

## 1.1. ВВОДНЫЕ ПОНЯТИЯ

Представим несколько вводных понятий из области биометрики и биометрических технологий.

**Биометрика** (Biometrics) — область знаний, представляющая методы измерения физических характеристик и формулирования персональных поведенческих черт человека и их использования для идентификации или аутентификации человека [Woo97].

Здесь следует отметить, что в отечественной научно-популярной литературе вместо термина «биометрика» (это соответствует прямому переводу английского слова «biometrics») используется термин «биометрия», который, в свою очередь, появился как перевод слова «biometry». Последний термин был введен в употребление Френсисом Гальтоном в 1889 г. в книге, посвященной *проблемам природной наследственности*. В дисциплину биометрии сегодня входят такие задачи, например, как измерение и обработка данных в биологических экспериментах, математическое моделирование в биологии, биологическое приложение теории планирования эксперимента и т. д. Разница между двумя приведенными выше терминами отчетливо прослеживается в публикациях [Лит22, Бар90, Зав98, Лео99].

В настоящей книге будет применяться термин биометрика, вытекающий из представлений [Woo97].

Собственно идентификация или аутентификация человека реализуется биометрической системой. В свою очередь, *Биометрическая система* — автоматизированная система, решающая задачи регистрации пользователей и их идентификацию (аутентификацию) и реализующая следующие функции:

- фиксацию биометрической выборки от конечного пользователя;
- извлечение биометрических данных из этой выборки;
- сравнение биометрических данных с одним или большим количеством эталонов;
- принятие решения о том, насколько хорошо соответствуют друг другу сравниваемые биометрические данные;
- формирование результата, например, информации о том, была или нет достигнута идентификация или аутентичность;






— принятие решения о повторении, окончании или видоизменении процесса идентификации (или аутентификации).

Методы и соответствующие им технические средства получения и использования биометрических данных человека в целях его идентификации (верификации, аутентификации или распознавания) определим как *биометрические технологии*. При этом в биометрических технологиях используются как *физические биометрические характеристики* человека — отпечатки пальцев, геометрия руки, изображения радужной оболочки и сетчатки глаза, голос, видео- и термоизображение лица, подпись (а в последнее время изображения и форма ушей, запах человека), так и *поведенческие биометрические характеристики*, например, манера работы на клавиатуре компьютера, динамика написания подписи, стиль и манера походки.

Некоторые примеры статических характеристик представлены в табл. 1.1.1. На практике используются как графические особенности (см. табл. 1.1.1 среднюю колонку), так и специфические особенности биометрических характеристик (см. правую колонку).

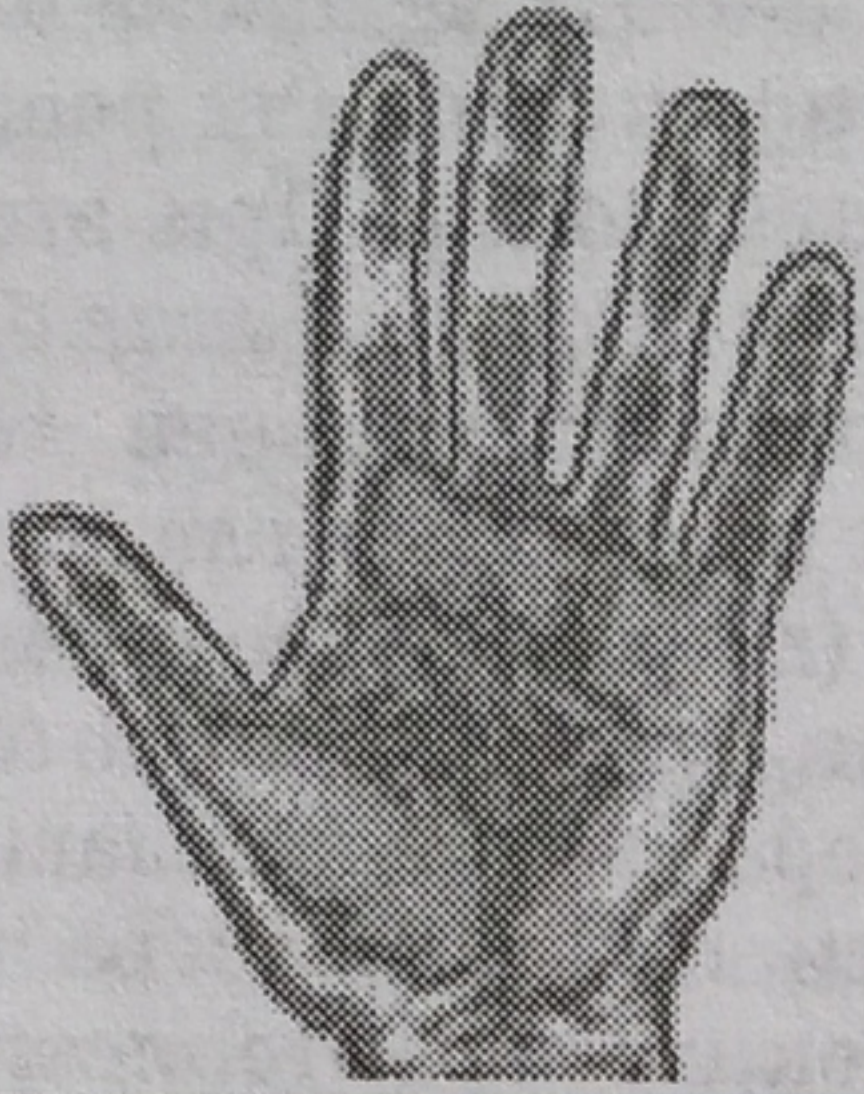

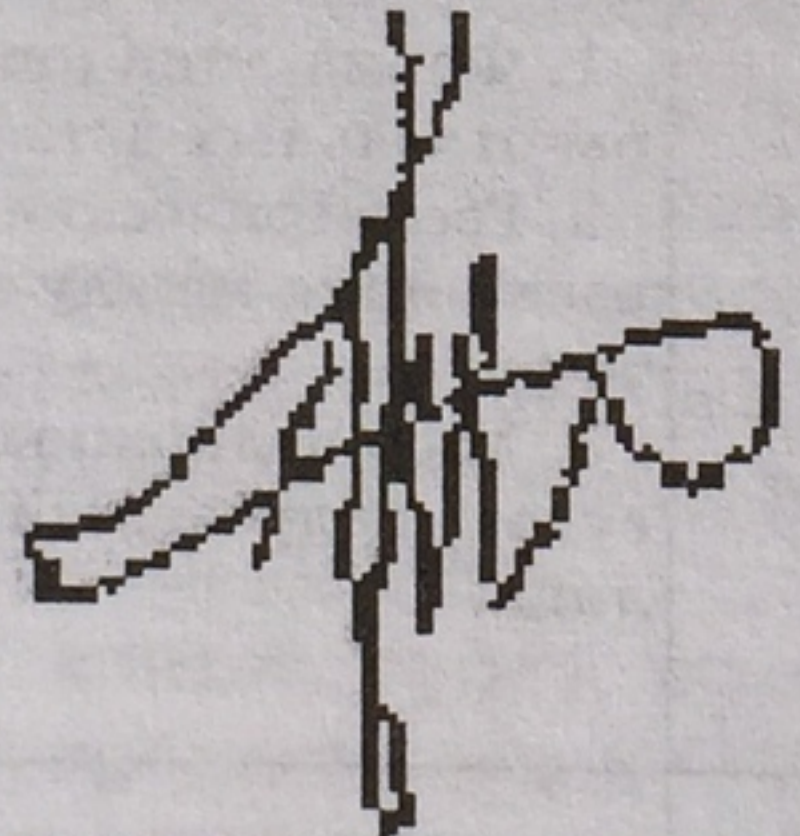
Т а б л и ц а 1.1.1

Примеры статических характеристик

| Характеристика | Пример  | Используемые особенности   |
|----------------|---|--|
| Лицо           |  | <ol style="list-style-type: none"><li>1. Форма лица (овал, форма и размеры отдельных деталей лица).</li><li>2. Геометрические параметры лица — расстояния между его определенными точками.</li><li>3. Узор подкожных кровеносных сосудов, получаемый как термограмма лица.</li></ol> |
| Глаз           |  | <ol style="list-style-type: none"><li>1. Рисунок радужной оболочки.</li><li>2. Узор кровеносных сосудов на сетчатке.</li></ol>   |
| Ухо            |  | <ol style="list-style-type: none"><li>1. Рисунок формы уха (контур и наклон, козелок и противокозелок, форма и прикрепление мочки и т. д.).</li><li>2. Параметры, выраженные как расстояния между определенными точками на ухе.</li></ol>  |



Продолжение табл. 1.1.1

| Характеристика   | Пример   | Используемые особенности   |
|------------------|--|--|
| Рука             |    | <ol style="list-style-type: none"> <li>1. Геометрия руки (ширина, длина, высота пальцев, расстояния между определенными точками).</li> <li>2. Неровности складок кожи на сгибах пальцев тыльной стороны кисти руки.</li> <li>3. Рисунок вен на тыльной стороне кисти руки, получаемый при инфракрасной подсветке.</li> <li>4. Узор на ладони.</li> </ol> |
| Отпечаток пальца |  | <ol style="list-style-type: none"> <li>1. Папиллярный узор как целостный образ.</li> <li>2. Параметры минуций (координаты, ориентация, тип).</li> </ol>  |
| Подпись          |  | <ol style="list-style-type: none"> <li>1. Подпись как двумерный бинарный образ.</li> <li>2. Подпись как функция двух координат.</li> <li>3. Динамика подписи (сила нажима и координата времени).</li> </ol>  |
| ...              | ...  | ...  |

Пример динамической (поведенческой) биометрической характеристики — динамики подписи — приведен на рис. 1.1.1.

Среди задач, которые можно решать с помощью биометрики, можно выделить, по крайней мере, три: идентификация (аутентификация) личности человека; оценка намерений и наклонностей индивидуума и медицинская диагностика состояния здоровья человека; контроль доступа. Примеры приложений, в которых используется идентификация личности человека, приведены на рис. 1.1.2.

Оценка эмоционального состояния человека, его чувств и намерений по отношению к окружающим важна в местах массового скопления людей. Идентифицировать эти состояния и намере-



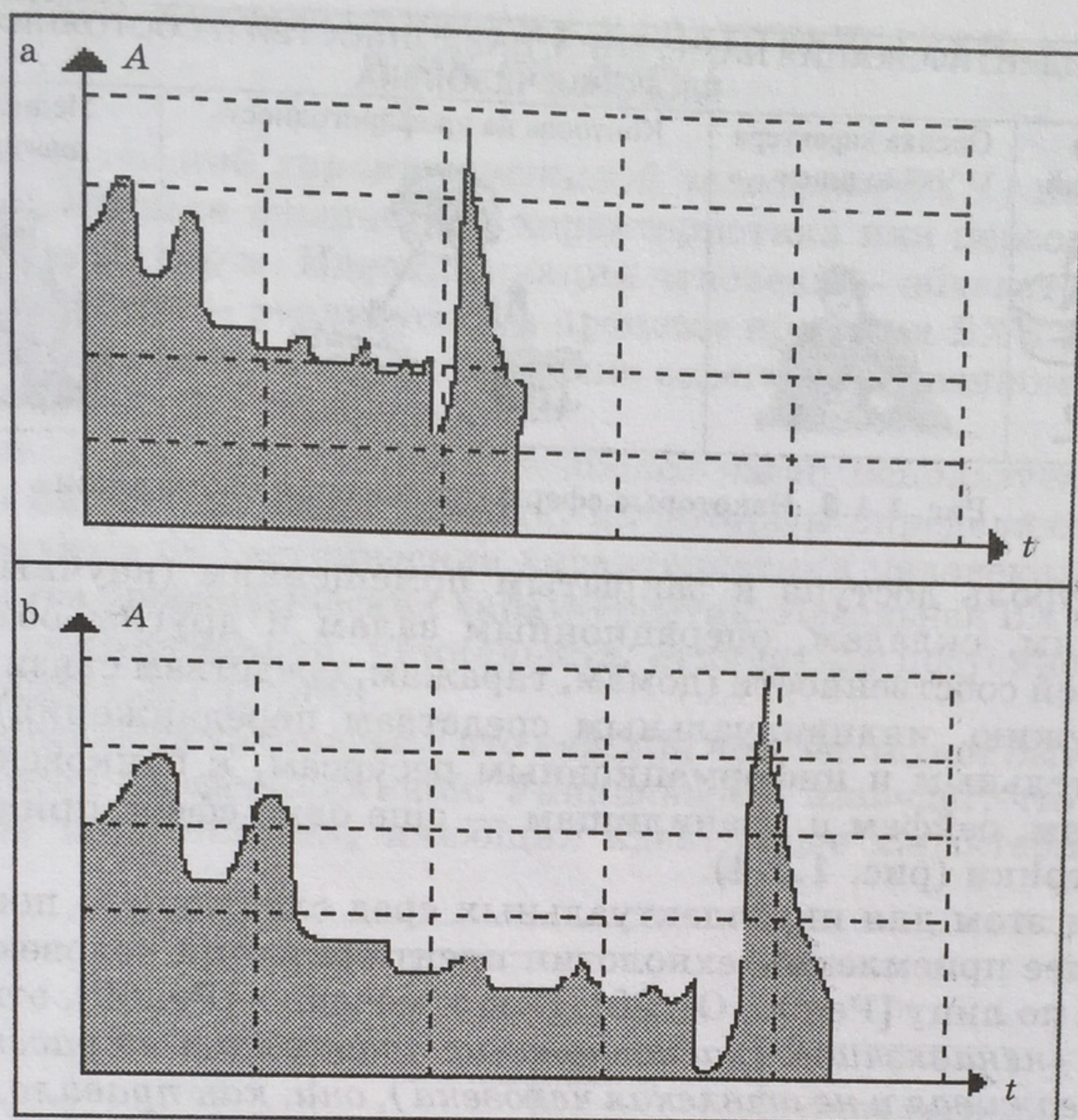


Рис. 1.1.1. Пример поведенческих биометрических характеристик: а — динамические характеристики  $A(t)$  (колебания пера и время написания подписи), принадлежащие автору подписи; б — те же характеристики для случая, когда подпись повторяет другое лицо

ния — значит предупредить возможные тяжелые последствия или непредвиденные ситуации.

Каково предназначение человека, какими он «наделен» болезнями или привычками, можно узнать по лицу человека (физиогномика), радужной оболочке глаза (иридоанализ, иридодиагностика), по почерку (графоанализ, графология) и голосу (рис. 1.1.3).

| ИДЕНТИФИКАЦИЯ ЛИЧНОСТИ ЧЕЛОВЕКА         |   |                        |
|---|---|------------------------|
| Криминалистика<br>и судебная экспертиза | Идентификация конечного<br>пользователя, клиента, гостя | Паспортный<br>контроль |
|   |   |                        |

Рис. 1.1.2. Примеры приложений идентификации (аутентификации) личности человека





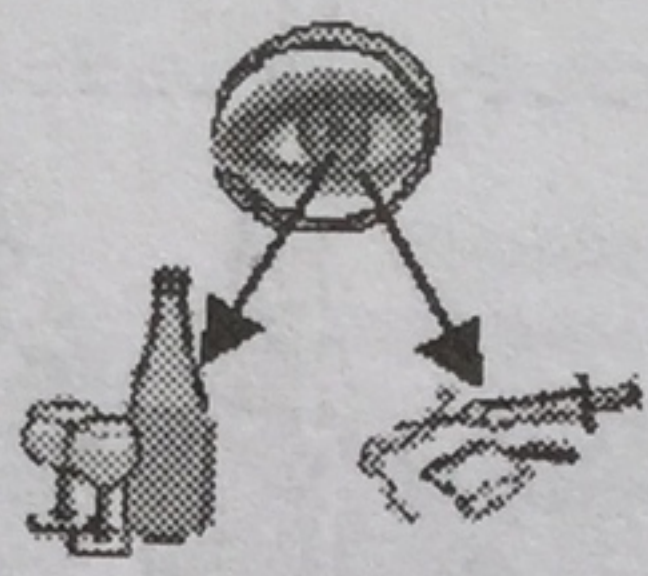

| ИДЕНТИФИКАЦИЯ НАМЕРЕНИЙ, НАКЛОННОСТЕЙ И СОСТОЯНИЯ<br>ЗДОРОВЬЯ ЧЕЛОВЕКА            |   |   |   |
|---|---|---|---|
| Оценка намерений  | Оценка характера по подписи   | Контроль на профпригодность   | Медицинская диагностика   |
|  |  |  |  |

Рис. 1.1.3. Некоторые сферы приложений биометрики

Контроль доступа к закрытым помещениям (научным лабораториям, складам, операционным залам и другим объектам), к личной собственности (домам, гаражам, средствам связи, личному оружию, индивидуальным средствам передвижения), к вычислительным и информационным ресурсам, к банковским терминалам, сейфам и хранилищам — еще одна сфера приложений биометрики (рис. 1.1.4).

При этом для интеллектуальных сред следующего поколения наиболее приемлемы технологии идентификации человека по голосу и по лицу [Реп00, Отс00]. Как отмечено в [Реп00], эти технологии «ненавязчивы (распознавание происходит на расстоянии, не задерживая и не отвлекая человека), они, как правило, пассивны (не требуют специальных действий со стороны человека), не ограничивают пользователя в свободе перемещений и к тому же относительно недороги. Но самое важное, по-видимому, то, что люди обычно узнают друг друга по голосу и по лицам довольно просто и, значит, не будут испытывать неудобств с подобными системами...»

Некоторым из этих проблем, в частности идентификации (верификации, аутентификации или распознаванию) человека на основе лица, и посвящена настоящая монография.

| КОНТРОЛЬ ДОСТУПА  |   |  |   |
|---|---|--|---|
| К закрытым помещениям   | К собственности и средствам передвижения  | К вычислительным и информационным ресурсам   | К банковским операциям  |
|  |  |  |  |

Рис. 1.1.4. Приложения биометрики, связанные с контролем доступа



## 1.2. БИОМЕТРИЧЕСКИЕ ХАРАКТЕРИСТИКИ И ИХ СВОЙСТВА

*Биометрической характеристикой человека (БХЧ)* называется его измеряемая физическая характеристика или персональная поведенческая черта. Идентификация человека — обладателя этой характеристики — реализуется в процессе проверки БХЧ на идентичность аналогичной характеристики зарегистрированному (санкционированному) пользователю.

На рис.1.2.1 представлены наиболее часто используемые источники биометрических данных, по которым определяются индивидуальные биометрические характеристики человека.

**Свойства биометрических характеристик.** Идеальная БХЧ должна быть универсальной, уникальной, отличаться постоянством и «собираемостью» [Jai00].

Универсальность означает, что каждый человек может быть представлен этой характеристикой. Уникальность означает, что не может быть двух человек, имеющих идентичные характеристики.



Рис. 1.2.1. Источники физических и поведенческих биометрических характеристик



Постоянство определяет независимость характеристики во времени, а собираемость — то, что характеристика может быть относительно просто и быстро получена от каждого индивидуума и так же может быть относительно просто и быстро детализирована.

Однако, как свидетельствует практика, реальные БХЧ не столь идеальны, что может существенно ограничивать их применение. В табл. 1.2.1 приведена экспертная оценка (в основном по материалам [Jai00]) свойств некоторых биометрических характеристик.

Как видно из табл. 1.2.1, ни одна из характеристик не имеет высоких оценок по всем параметрам, хотя, стоит заметить, выбор конкретной БХЧ определяется не столько этими оценками, сколько условиями применения систем идентификации — их назначением, степенью ответственности при принятии решений, степенью защиты от несанкционированного доступа к системе (степенью защиты от взлома системы) и, наконец, стоимостью самой системы.

**Точность и стоимость биометрических систем.** Примерные соотношения различных биометрических технологий по затратам на их реализацию и по достижимой с их использованием точности идентификации людей представлены схематически на рис. 1.2.2 [Пан00].

Затраты здесь складываются из затрат на закупку сканеров (датчиков) биометрической информации, разработку и реализацию собственно биометрической системы (проект + компьютер-

Экспертная оценка свойств БХЧ

Таблица 1.2.1

| Характеристика  | Универсальность | Уникальность | Постоянство | Собираемость |
|---|-----------------|--------------|-------------|--------------|
| Видеообраз лица   | +++             | +            | ++          | +++          |
| Термограмма лица  | +++             | +++          | +           | +++          |
| Отпечаток пальца  | ++              | +++          | +++         | ++           |
| Геометрия руки  | ++              | ++           | ++          | +++          |
| Радужная оболочка глаз                                      | +++             | +++          | +++         | ++           |
| Сетчатка  | +++             | +++          | ++          | +            |
| Подпись   | +               | +            | +           | +++          |
| Голос   | ++              | +            | +           | ++           |
| Отпечаток губ   | +++             | +++          | ++          | +            |
| Особенности ушной раковины                                  | ++              | ++           | ++          | ++           |
|   |                 |              |             |              |
| Динамика письма   | +++             | +++          | +           | +++          |
| Походка   | +++             | ++           | +           | +            |
| Здесь (+++) — высокая оценка; (++) — средняя; (+) — низкая. |                 |              |             |              |



ная система + математическое обеспечение + техническое обслуживание), а также затрат на ее эксплуатацию.

Как следует из рис. 1.2.2, очевидны две противоречивые тенденции: первая (отмечена наклонной стрелкой в направлении: *голос* → *лицо* → *палец* → *сетчатка* → *радужная оболочка*) как бы свидетельствует о том, что точность биометрических технологий прямо связана с затратами на их реализацию. Вторая (прослеживается в направлении: *палец* → *подпись* (*уши*) → *рука*) говорит о том, что дорогие технологии не всегда самые точные [Пан00]. Так от чего же зависит выбор той или иной биометрической технологии для конкретных приложений?

Обратим внимание на технологию, использующую отпечатки пальцев. На рис. 1.2.2 эта технология оценена как имеющая средний показатель затраты/точность. Однако по оценке международной группы по биометрике (International Biometric Group — BIG, [Bio\*\*]) данная технология была в 1999 г. лидером продаж среди прочих и занимала около 40 % рынка биометрических технологий.

Общая картина продаж средств биометрических технологий (по данным BIG) представлена в табл. 1.2.2.

Лидерство технологии, использующей отпечатки пальцев, на рынке продаж объясняется следующим. Во-первых, сказывается опыт длительного (исторического) применения отпечатков пальцев в криминалистике и связанное с этим утверждение международных стандартов на отпечатки пальцев (+ базы отпечатков), во-вторых, наличие широкой номенклатуры разработанных и внедренных в использование сканеров отпечатков пальцев, в-третьих, развитая система сканеров, встроенных в компьютерное оборудование (в мыши, клавиатуру, ноутбуки), в-четвертых, существование готовых к приме-

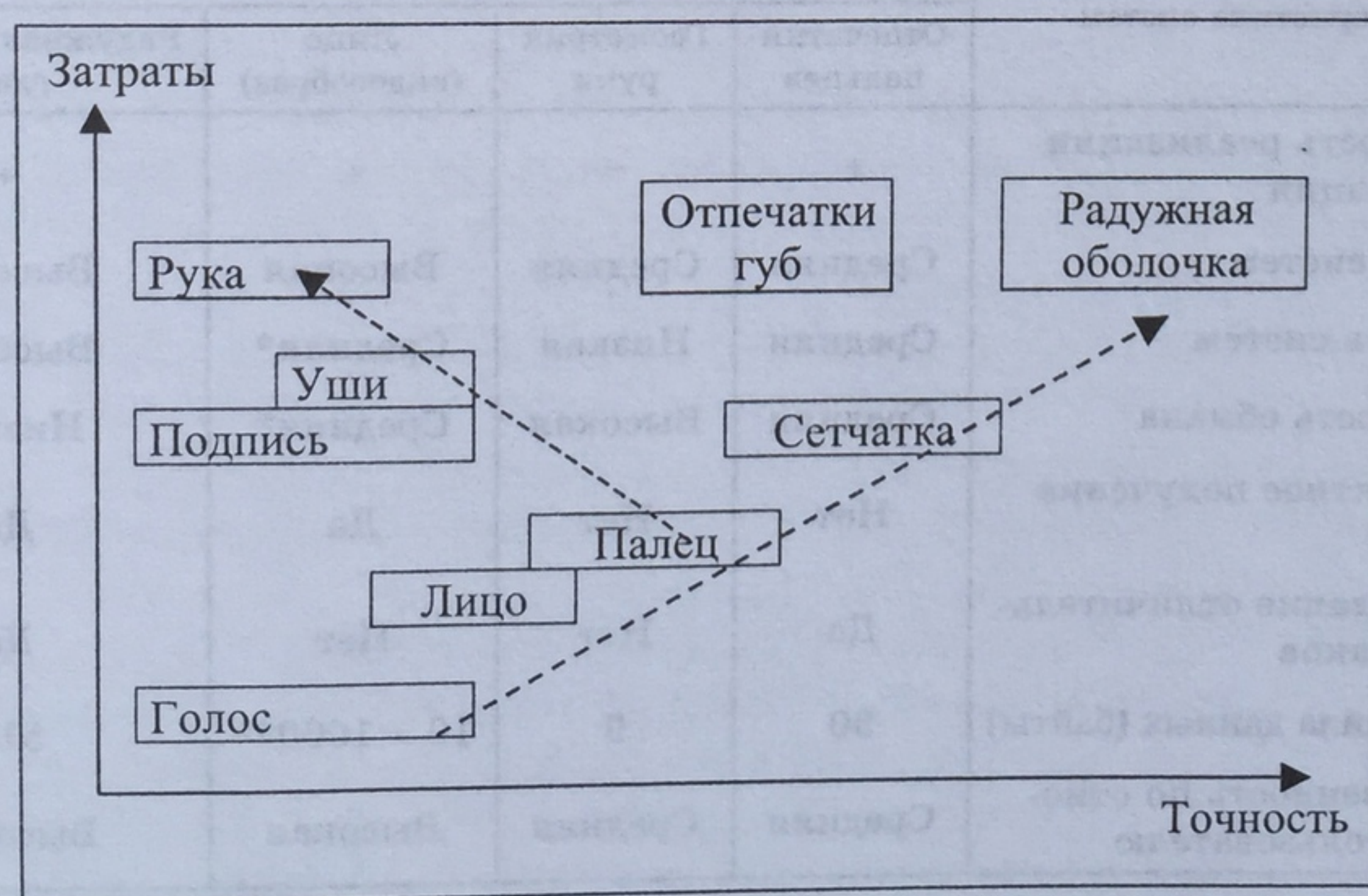


Рис. 1.2.2. Затраты и точность различных биометрических технологий [Пан00]



Т а б л и ц а 1.2.2

Продажа средств биометрических технологий в 1999 году

| Средства биометрических технологий                         | Объем продаж, % |
|--|-----------------|
| Терминалы и устройства идентификации по отпечаткам пальцев | 34              |
| Устройства распознавания по геометрии руки                 | 26              |
| Системы распознавания по лицам                             | 15              |
| Сканеры и системы идентификации по радужной оболочке глаза | 9               |
| Сканеры и системы идентификации по сетчатке глаза          | 2               |
| Системы идентификации по голосу                            | 11              |
| Системы верификации по подписи                             | 3               |

нению законченных устройств и терминалов идентификации отпечатков, которые продаются в комплекте с системами AFIS, и, наконец, в-пятых, наличие подробно описанных и проверенных практикой алгоритмов идентификации отпечатков пальцев.

Представим теперь оценки биометрических технологий и основанных на них систем с точки зрения пользователей этих систем [Bio\*\*]. При этом отметим такие характеристики, как:

- возможность решения задачи идентификации (а не аутентификации);
- сложность реализации систем идентификации;
- достигаемую при этом точность идентификации;

Т а б л и ц а 1.2.3

Оценки четырех биометрических технологий

| Характеристика систем  | Технология        |                |                   |                         |
|--|-------------------|----------------|-------------------|-------------------------|
|  | Отпечатки пальцев | Геометрия руки | Лицо (видеообраз) | Радужная оболочка глаза |
| Возможность реализации идентификации   | ±                 | —              | +                 | +                       |
| Точность систем  | Средняя           | Средняя        | Высокая           | Высокая                 |
| Сложность систем   | Средняя           | Низкая         | Средняя*          | Высокая                 |
| Возможность обмана   | Средняя           | Высокая        | Средняя*          | Низкая                  |
| Бесконтактное получение данных   | Нет               | Нет            | Да                | Да                      |
| Использование отличительных признаков  | Да                | Нет            | Нет               | Да                      |
| Размер файла данных (байты)  | 90                | 9              | 10 + 1000**       | 512                     |
| Дружелюбность по отношению к пользователю  | Средняя           | Средняя        | Высокая           | Высокая                 |
| * — в зависимости от назначения системы;   |                   |                |                   |                         |
| ** — в зависимости от назначения системы, объема базы данных и способа редукции исходных данных. |                   |                |                   |                         |



Адреса сайтов, посвященных биометрике

Таблица 1.2.4

| Технология        | Адрес страниц  |
|-------------------|--|
| Отпечаток пальца  | www.biometricaccess.com, www.identix.com<br>www.biometricid.com, www.abio.com, www.rjm-jena.de,<br>www.digitalpersona.com, www.startek.com.tw,<br>www.iosoftware.com, www.identicator.com. |
| Геометрия руки    | www.recogsys.com   |
| Лицо (видеообраз) | www.faceit.com, www.miros.com, www.neurodynamics.com,<br>www.viisage.com.  |
| Радужная оболочка | www.iriscan.com, www.sensar.com.   |

- степень возможности обмана системы (несанкционированного доступа);
- возможность бесконтактного получения биометрических данных;
- размер файла эталона биометрической характеристики;
- дружелюбность по отношению к пользователю.

Эти оценки для четырех наиболее популярных биометрических технологий сведены в табл. 1.2.3.

Как видно из табл. 1.2.3, две последние технологии могут вызывать повышенный интерес для практических приложений биометрики.

Новости о биометрике и ее приложениях, а также демо-версии систем распознавания биометрических образов (например, VeriFinger, TrueFace и др.), можно найти на ряде сайтов (табл. 1.2.4).



## 2. ОБЗОР НЕКОТОРЫХ БИОМЕТРИЧЕСКИХ ТЕХНОЛОГИЙ

### 2.1. ТЕХНОЛОГИЯ ОТПЕЧАТКОВ ПАЛЬЦЕВ. ИСТОРИЯ ПРИМЕНЕНИЯ

В качестве одного из аргументов применения технологии на основе отпечатков пальцев по сравнению, например, с технологией распознавания человека на основе его лица авторы [Jai97a], известные своими работами в области AFIS, приводят фотографии, на которых представлено несколько сцен в уличном кафе. Все персоны на фотографиях — официант, посетитель кафе, прохожий, роликбежец и т. д. — представляют как бы разных людей (они имеют разные прически и камуфляж — повязки и шапочки на голове, очки, усы...), и система Face Recognition приняла бы их всех за разных людей. Однако все персоны на фотографиях — один и тот же человек! Очевидно, отмечают авторы работы [Jai97a], что с отпечатками пальцев невозможно сделать ничего подобного.

Отпечатки пальцев для распознавания личности применялись уже в Древнем Египте, Ассирии, Китае и Японии [Поо98]. В этих странах для удостоверения личности в момент заключения различных сделок, использовали отпечаток большого пальца. К пионерам дактилоскопии в XIX в. можно отнести профессора Я.-Э. Пуркинью, врача Генри Фолдса, писателя Марка Твена, антрополога Френсиса Гальтона [Поо98].

Первый научный трактат о папиллярных узорах на пальцах написал в 1823 г. профессор Бреславльского университета Я.-Э. Пуркинью. Он выделил и описал узоры (рис. 2.1.1), которые используются в дактилоскопии до настоящего времени.

Врач Генри Фолдс в 1824 г. представил чернильный отпечаток пальца как средство идентификации человека. Он же первым использовал отпечаток пальца в практике идентификации преступника (отпечатки пальцев преступник оставил на бутылке изпод алкогольного напитка).

Марк Твен в 1883 г. написал книгу «Жизнь на Миссисипи», в которой убийца был идентифицирован на основе собственного отпечатка пальца. В позднейших своих изданиях Марк Твен также описывал процессы поиска преступников с помощью отпечатков пальцев.

Английский антрополог Френсис Гальтон в 1892 г. публикует книгу, в которой обращает внимание на уникальность характеристик отпечатков пальцев каждого человека и их постоянство во времени. Он также представил первую систему признаков, по которой можно классифицировать отпечатки пальцев.

Главное, что вытекает из публикаций Ф. Гальтона и других пионеров дактилоскопии: в основе метода идентификации лич-

нос  
уни  
дям  
Б  
ный  
Б  
цией  
том  
2.1.  
В  
типо  
каль  
К  
деть  
зом:  
тип  
Г.  
•  
•  
чатк  
•  
кой-л





Рис. 2.1.1. Некоторые типы папиллярных узоров на пальцах:

- 1 – 4 — узоры типа «петля» (соответственно левая, правая, центральная, двойная),  
 5 и 6 — узоры типа «дельты» или дуги (соответственно простая и острая), а 7 и  
 8 — узоры типа «спираль» (соответственно центральная и смешанная)

ности человека на основе отпечатков пальцев лежит факт их уникальности. Отпечатки пальцев, принадлежащие разным людям, не могут быть идентичными.

В XX в. этот метод получил официальное признание как научный на Международном полицейском конгрессе (Монако, 1914 г.).

В советской России начала XX в. дактилоскопической регистрацией занимался П. С. Семеновский — первый заведующий кабинетом судебной экспертизы и регистрационным бюро Центроорозыска.

### 2.1.1. ИНФОРМАЦИОННЫЕ ПРИЗНАКИ ОТПЕЧАТКОВ ПАЛЬЦЕВ

В каждом отпечатке пальца можно определить признаки двух типов, используемые при их идентификации: глобальные и локальные [Rat96, Cap99].

К *первому типу* относятся признаки, которые можно разглядеть и/или различить на отпечатках пальцев невооруженным глазом: папиллярные узоры; область образа; ядро; пункты «дельта»; тип папиллярных линий и их число.

Глобальные признаки интерпретируются следующим образом:

- *узоры папиллярные*, о них мы вспоминали в разд. 1.2;
- *область образа* — некоторый выделенный фрагмент отпечатка, в котором локализованы все признаки;
- *ядро* — пункт, локализованный в середине отпечатка или в какой-либо выделенной области;



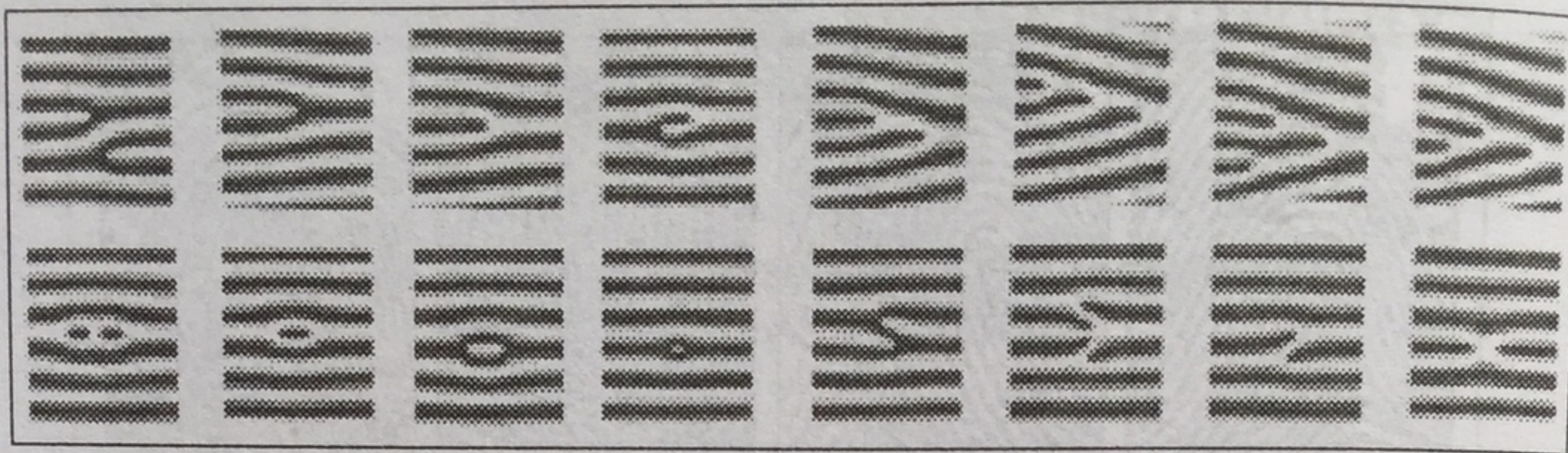


Рис. 2.1.2. Примеры микроузоров на основе двух типов минуций

- *пункт «дельта»* — начальная точка, то место, в котором происходит разделение бороздки папиллярной линии или место, в котором соединяются две бороздки, или, наконец, очень короткая (доходящая до точки) бороздка;

- *тип линии* — две наибольшие линии, которые начинаются как параллельные, а затем расходятся и огибают всю область образа;

- *счетчик линий* — число линий (бороздок) на области образа или между ядром и пунктом дельта.

Ко второму типу относятся локальные признаки, называемые *минуциями*: это уникальные для каждого отпечатка признаки, определяющие пункты изменения структуры папиллярных линий (например, разрыв, окончание, раздвоение и т. д.), ориентацию папиллярных линий и координаты  $X$  и  $Y$  в этих пунктах. Каждый отпечаток может быть представлен 50–70 минуциями. Фирма Optel [Bio\*\*] предоставила снимок различных микроузоров, которые могут образовываться из двух типов минуций — окончания и раздвоения гребней папиллярных линий (рис. 2.1.2).

Практика показывает, что в отпечатках пальцев разных людей могут встречаться идентичные глобальные признаки, однако совершенно невозможно наличие в этих же отпечатках одинаковых микроузоров минуций. Поэтому глобальные признаки используют при классификации (разделении на части) очень больших баз отпечатков пальцев и в иерархических процедурах идентификации отпечатков на этапах их предварительной классификации [Rat96, Cap99]. В последующем, на втором этапе распознавания отпечатка пальца, используют уже его локальные признаки.

В процедурах аутентификации глобальные признаки могут быть достаточными для точного сравнения двух отпечатков пальцев.

### 2.1.2. СТАНДАРТЫ НА ОТПЕЧАТКИ ПАЛЬЦЕВ

К концу 70-х гг. XX в. в архивах федеральных служб США находилось более 200 млн отпечатков пальцев, а в специальные отделы ФБР ежедневно приходило до 30 тыс. запросов на их идентификацию [Tou74]. Оценить сегодня суммарный объем баз данных отпе-

чатков  
жается  
Соз  
ков па  
связа  
разны  
кации  
все эт  
чатки

На  
и ФБР  
рые па  
ципиа  
процед  
способ  
жений  
(или р  
ния яр  
метров  
тами о

В ц  
образу

- ка
- скомпр
- об
- об
- ма
- ос

дарт Ф

Для  
тых с о  
чатков  
печатк  
System  
одного  
привед  
составл



Рис. 2.1.3



чатков пальцев не представляется возможным — их число приближается, скорее всего, к числу взрослых людей на земном шаре.

Создание и пополнение баз данных, идентификация отпечатков пальцев в национальных и транснациональных масштабах и связанное с этим выполнение операций с данными, полученными разными способами, использование различных систем идентификации (Automated Fingerprint Identification Systems, или AFIS) — все это потребовало установления стандартов на исходные отпечатки пальцев.

На сегодняшний день в основном используются стандарты ANSI и ФБР США [Ame\*\*, Bio\*\*, Fin00], в которых определены некоторые параметры изображений отпечатков пальцев, имеющие принципиальное значение в стандартизации построения баз данных и процедур проверки качества распознавания отпечатков, в том числе способы получения отпечатков, их представления в форме изображений, качество их как цифровых изображений — разрешение (или размер изображения отпечатка), число уровней представления яркости изображения, максимальные значения границ параметров искажений, например, поворота и т. д. Этими же стандартами определяются основные типы минуций.

В целом стандартами установлены следующие требования к образу отпечатка:

- каждый образ отпечатка представляется в формате TIF в несжатом виде;
- образ должен иметь разрешение не менее 500 dpi;
- образ должен быть полутоновым с 256 уровнями яркости;
- максимальный угол поворота образа от вертикали  $\pm 15^\circ$ ;
- основные типы минуций — окончание и раздвоение (стандарт ФБР США).

Для ряда случаев стандарты устанавливают число образов, снятых с одного пальца и составляющих отдельный класс базы отпечатков пальцев. В результате некоторые базы содержат до 10 отпечатков с каждого пальца. Например, база фирмы Biometric System Lab [Fin00] в одном классе содержит по восемь отпечатков одного и того же пальца; четыре отпечатка одного и того же пальца приведены на рис. 2.1.3. В исходной базе размер каждого образа составляет  $256 \times 256$  пикселей.



Рис. 2.1.3. Изображения отпечатков одного и того же пальца, составляющих один класс



Наличие в базе данных нескольких отпечатков одного и того же пальца призвано улучшить качество распознавания в режимах «on-line», поскольку вся совокупность образов в классе «охватывает» несколько различных вариантов изменений характеристик отпечатка.

На рис. 2.1.3 видно, что образы отличаются друг от друга сдвигом (перемещением центра отпечатка вверх и вправо) и небольшим, но заметным поворотом образов влево и вправо. Масштаб всей совокупности представленных выше образов оставался практически неизменным, поскольку, очевидно, все они были получены с одного и того же устройства сканирования отпечатков.

Если оценивать подобие представленных отпечатков как бы «сверху», то можно утверждать, что все они действительно получены с одного пальца.

### 2.1.3. УСТРОЙСТВА ДЛЯ ПОЛУЧЕНИЯ ОТПЕЧАТКОВ ПАЛЬЦЕВ

Выше отмечалось наличие широкой номенклатуры сканеров для получения отпечатков пальцев, в том числе встроенных в компьютерное оборудование (в мыши, клавиатуру, ноутбуки), а также готовых к применению законченных устройств и терминалов идентификации отпечатков, которые продаются в комплекте с системами AFIS.

Ниже представим некоторые из этих устройств, спецификации на которые можно найти в [Mat00, Ots00, Bio\*\*, Fin00], а также на сайтах фирм — производителей этих устройств. Адреса сайтов были приведены в табл. 1.2.4 в первом разделе.

Прежде всего отметим, что независимо от устройства, с помощью которого получен отпечаток пальца (емкостной датчик, элемент, чувствительный к температуре папиллярных линий, или оптический сканер), сам процесс их получения выглядит не сложнее, чем представлено на рис. 2.1.4.

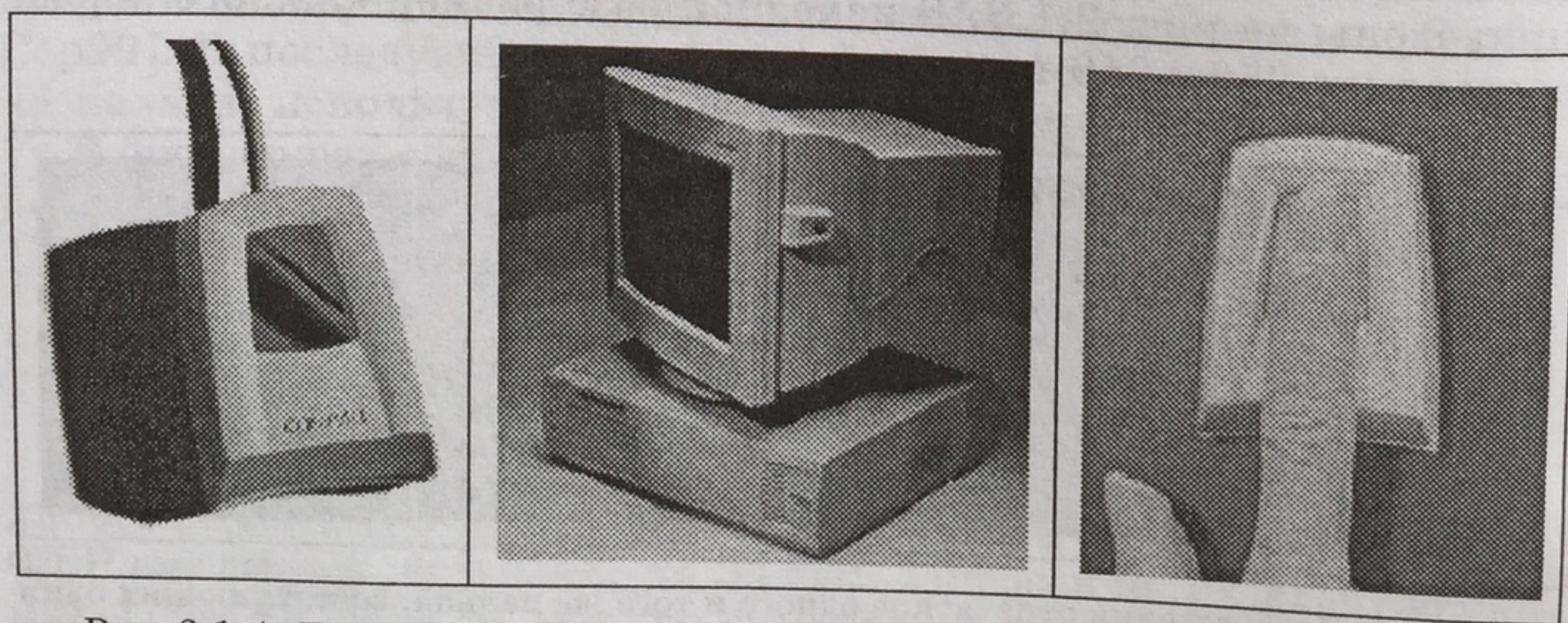


Рис. 2.1.4. Датчик, его крепление на мониторе и снятие отпечатка пальца

Замети  
ные устро  
мые к ко  
USB-интер  
мы COM  
ный на ри  
ство FCSw  
на рис. 2.  
такие инт  
лает их у  
простыми  
[Mat00, Ot  
Компа  
Technolog  
Inc.» ра  
U-Match™  
темах кон  
компьютер  
Достоин  
ватель авт  
пользован  
него само  
се использ  
ютерной м  
тиска боль  
Примен  
вателя об  
сетевой ил  
лее важно  
когда не с  
здает 500-  
быть реко  
мым обесп  
системы ид

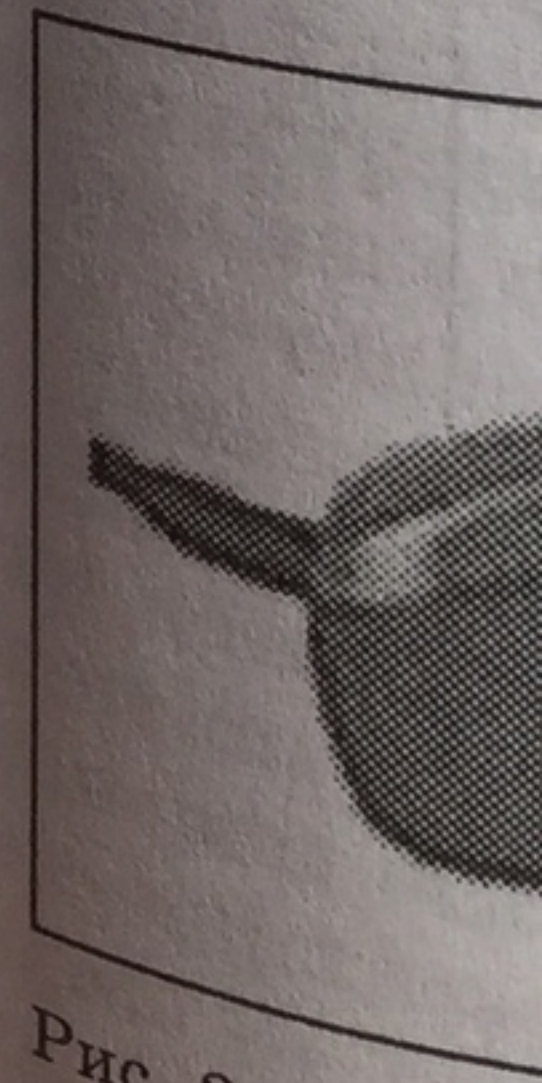


Рис. 2.1.6. М  
встроенным у



Заметим, что все отдельные устройства, подключаемые к компьютерам, имеют USB-интерфейс. Датчик фирмы COMPAQ, представленный на рис. 2.1.4, и устройство FCSweep02, показанное на рис. 2.1.5, имеют именно такие интерфейсы, что делает их универсальными и простыми в эксплуатации [Mat00, Otc00, Bio\*\*, Fin00].

Компанией «BioLink Technologies International, Inc.» разработана мышь

U-Match™ BioLink Mouse, предназначенная для использования в системах контроля доступа конечного пользователя к персональному компьютеру или сети компьютеров [см. [webmaster@biolinkusa.com](mailto:webmaster@biolinkusa.com)].

Достоинство применения такой мыши состоит в том, что пользователь автоматически верифицируется при каждом подходе к использованию компьютера, а также может быть (незаметно для него самого) подвергнут дополнительной верификации в процессе использования компьютера. Внешний вид стандартной компьютерной мыши, содержащей встроенный оптический сканер отпечатка большого пальца, представлен на рис. 2.1.6.

Применение этой мыши для идентификации конечного пользователя обеспечивает 100%-ную аутентификацию в локальной сетевой или межсетевой среде через BioLink-технологии. Наиболее важно то, что собственно отпечаток пальца пользователя никогда не сохраняется. Взамен этого мышь U-Match™ BioLink создает 500-байтовый зашифрованный эталон, который не может быть реконструирован в отпечаток пальца пользователя. Тем самым обеспечивается дополнительная защита от угрозы обмана системы идентификации или специфического несанкционированного доступа к ней.

Компания «BioLink Technologies International, Inc.» отмечает следующие достоинства своих систем идентификации конечного пользователя:

- размер сканируемого образа пальца: 284 × 400 пикселей;
- время регистрации 0,13 с;
- размер эталона отпечатка пальца 500 байтов.

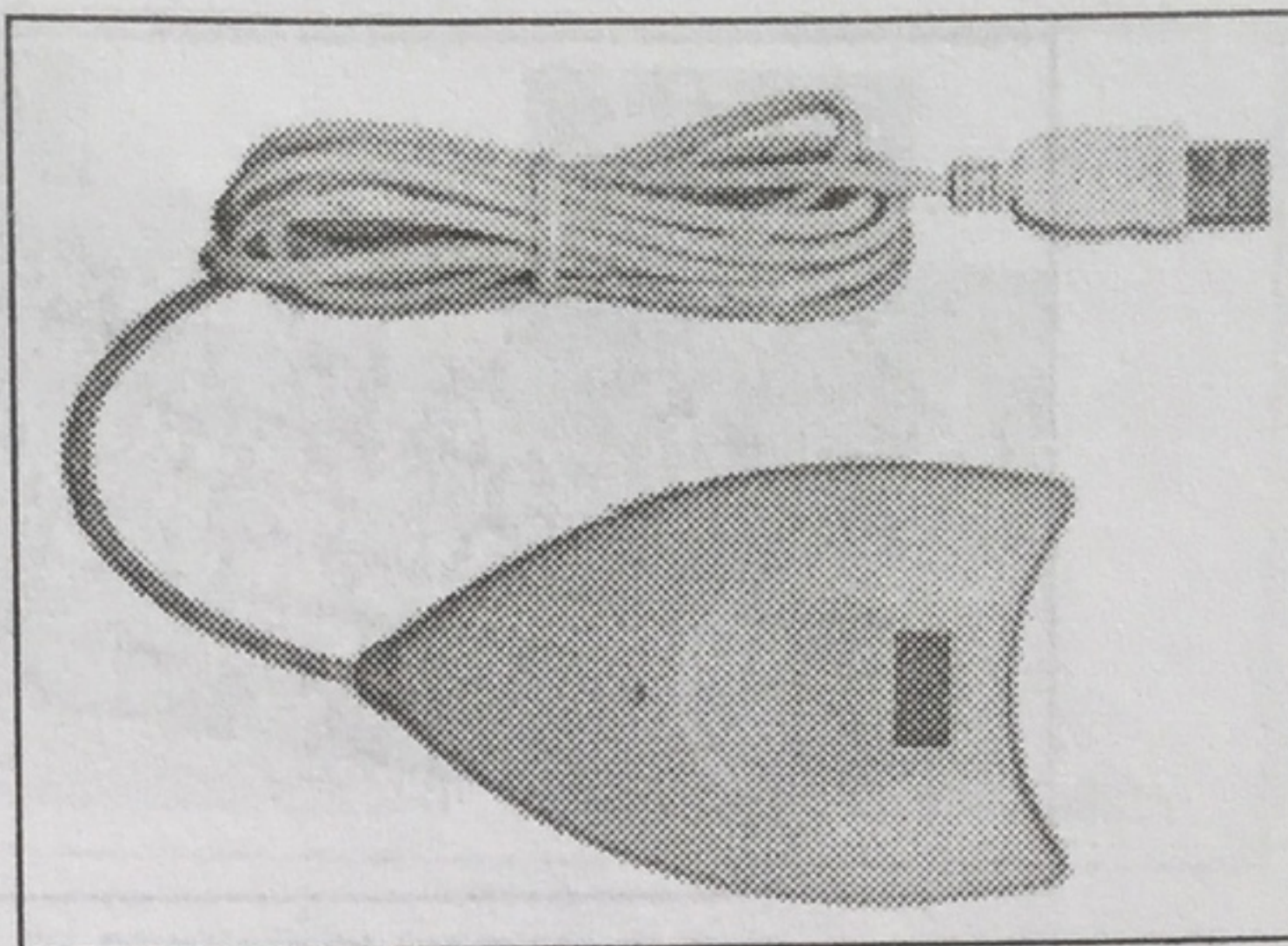


Рис. 2.1.5. Устройство считывания отпечатков пальца FCSweep02

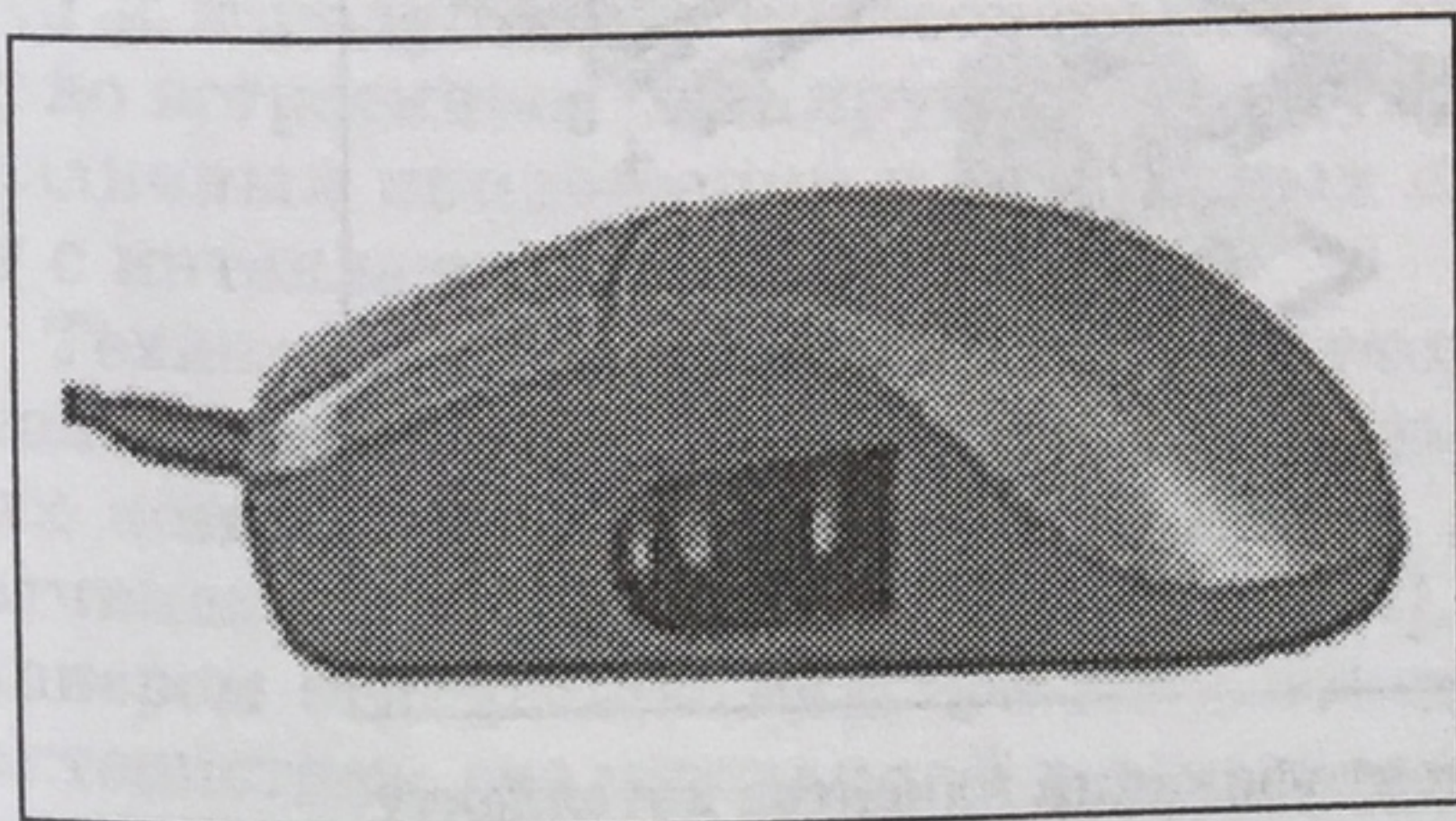


Рис. 2.1.6. Мышь U-Match™ BioLink Mouse со встроенным устройством сканирования отпечатка пальца



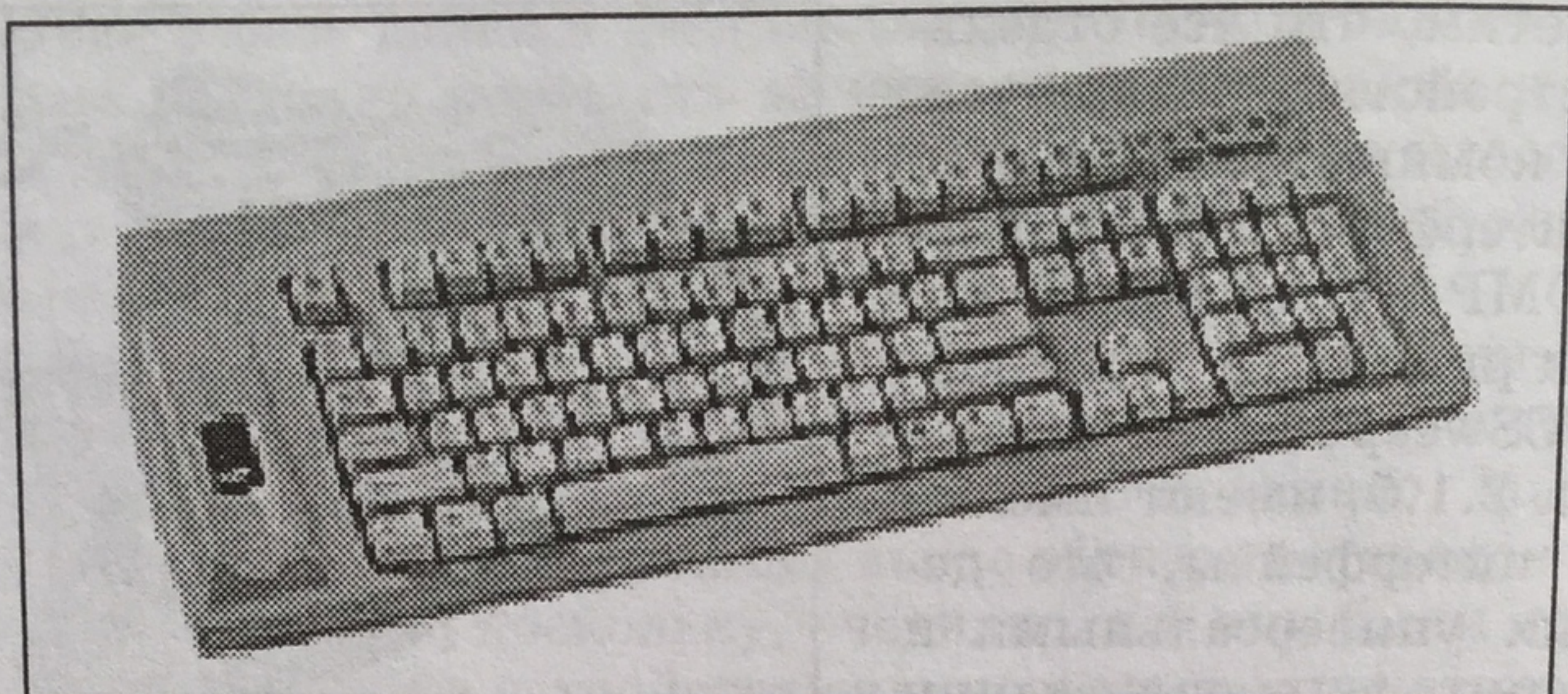


Рис. 2.1.7. Клавиатура со встроенным датчиком — сканером отпечатка пальца

Фирма «Keytronic» предлагает своим клиентам клавиатуру со встроенным датчиком — сканером отпечатка пальца. Внешний вид такой клавиатуры показан на рис. 2.1.7 [Мат00, Отс00, Bio\*\*, Fin00].

На современном этапе развития биометрических технологий некоторыми фирмами созданы и предлагаются на рынке целые системы контроля доступа пользователей к интеллектуальным средам. Такие системы включают весь состав компьютерного оборудования вместе с датчиками и оснащены как сканерами отпечатков пальцев прямого доступа пользователя, так и сканерами для мобильных пользователей (на базе смарт-карт). Внешний вид системы BioLogon 2.0 показан на рис. 2.1.8. Система предназна-

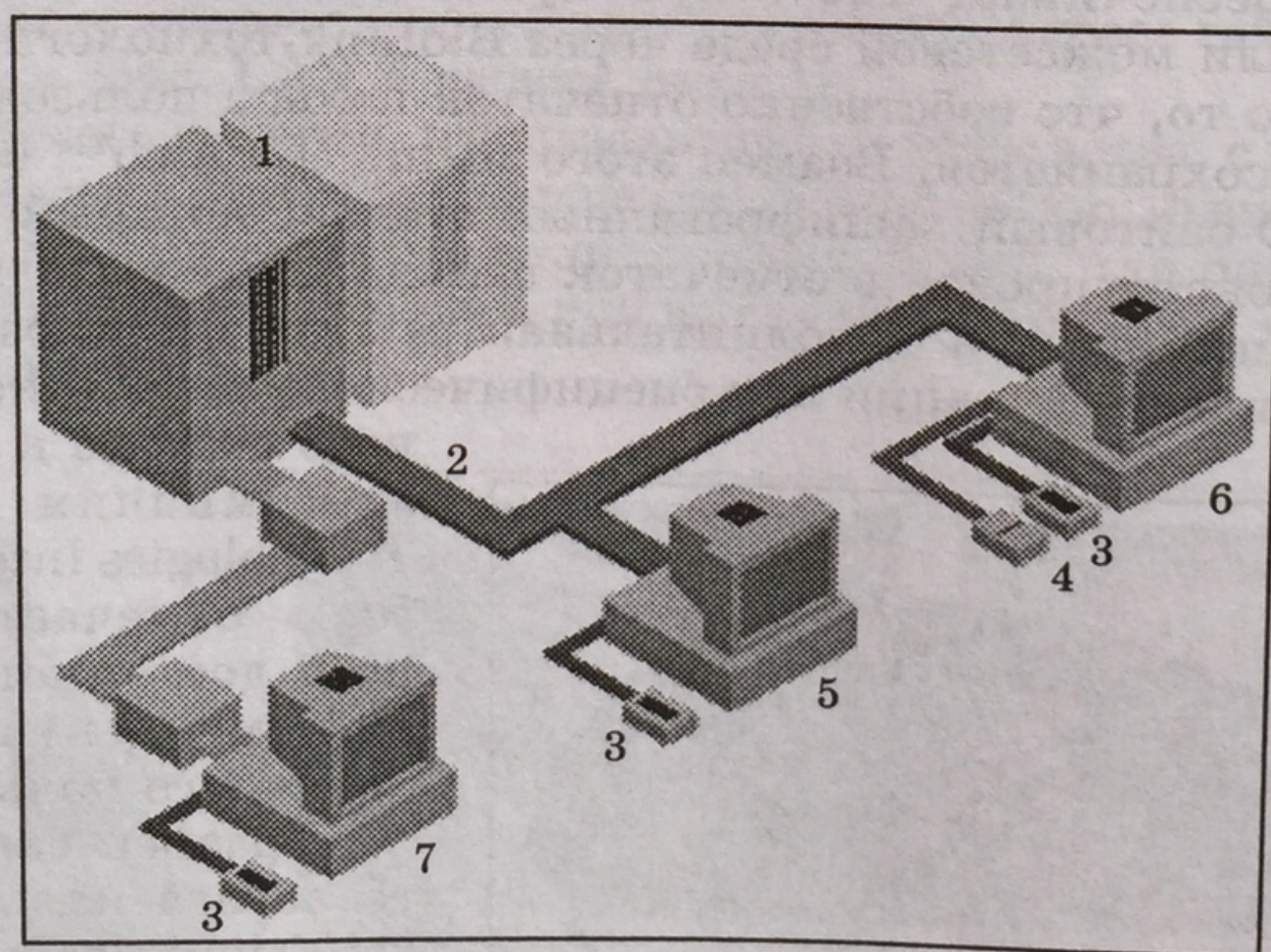


Рис. 2.1.8. Система аутентификации клиентов интеллектуальных сред по отпечаткам:

- 1 — сервер; 2 — шина интернет; 3 — датчик прямого действия; 4 — датчик для смарт-карты; 5 — место пользователя (клиента); 6 — место мобильного клиента; 7 — место удаленного клиента



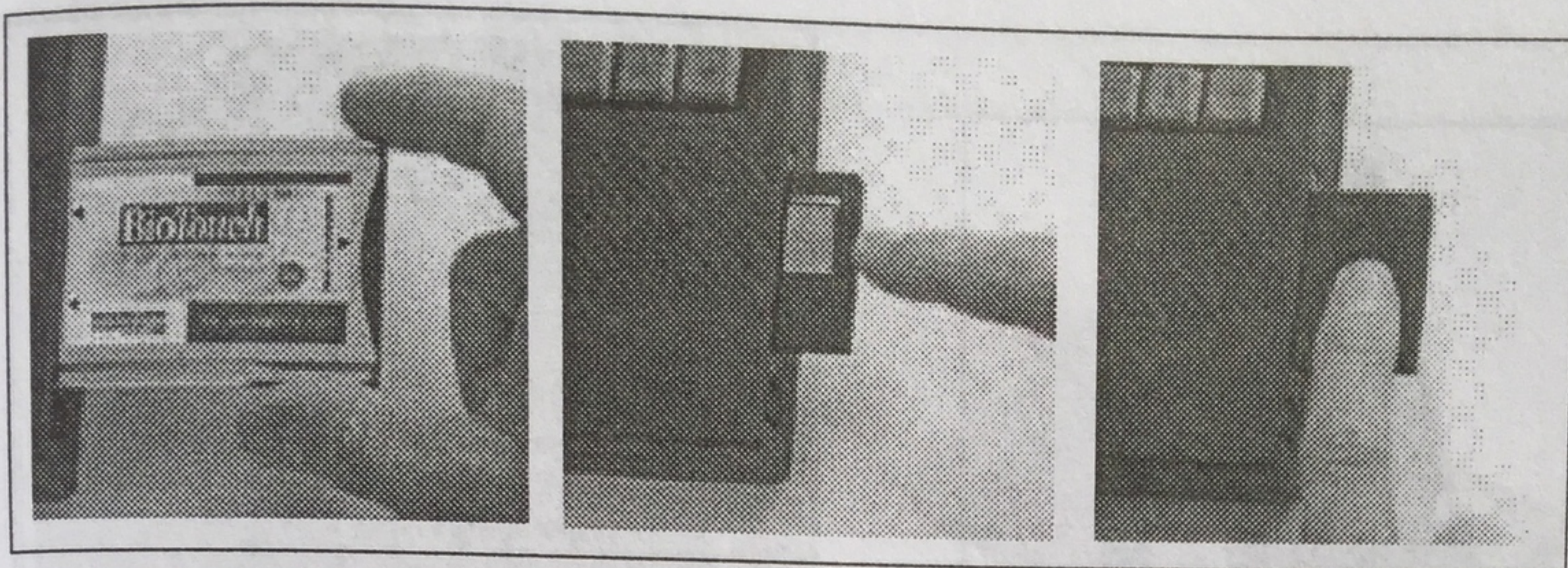


Рис. 2.1.9. Последовательность работы со сканером пальца в виде переносимой карты

чена для аутентификации клиентов (в том числе удаленных и мобильных), имеющих доступ к интеллектуальным средам широкого назначения.

Пользователи, не имеющие в базах данных системы своих отпечатков, хранят их на смарт-картах. В этом случае аутентификация клиента осуществляется путем сравнения информации (закодированного отпечатка пальца) на смарт-карте с «живым» отпечатком, полученным с помощью датчика прямого доступа.

Кроме того, фирмами, производящими оборудование в сфере биометрических технологий, предлагаются переносные датчики, встроенные в специальные карты и предназначенные для «мобильных» пользователей. Использование такой карты и получение отпечатка показано на рис. 2.1.9 [Mat00, [www.east-shore.com/tech.html](http://www.east-shore.com/tech.html)].

Примером автономного устройства, передающим на расстояние информацию об отпечатках пальца его владельца, или, например, индивидуального пилота дальнего управления (телевизором, гаражом, замковыми механизмами дверей автомобиля), или, наконец, сотового телефона с автоматической аутентификацией пользователя, является переносное устройство DERMALOG с датчиком отпечатков пальцев, показанное на рис. 2.1.10 [info@dermalog.de].

Оно используется в биометрических системах для идентификации и верификации пользователя по отпечаткам пальцев и оснащено встроенным сканирующим модулем. Идеально подходит для различных приложений в мобильных системах связи пользователя с интеллектуальными средами.

Технология распознавания отпечатков пальцев для защиты доступа к переносному компьютеру использована при оснащении таких компьютеров системами AFIS и встроенными в компьютер датчиками отпечатков пальцев [Com00]. Компьютер с встроенным сканером отпечатков пальцев показан на рис. 2.1.11 [Com00]. Характеристики реализованной в нем системы AFIS следующие: платформа — Laptop или Notebook с процессором Pentium. Сканируемое изображение пальца: 500 dpi, 256 уровней яркости. Размер файла для каждого эталона отпечатка пальца — от 256 байт до 4 кбайт





Рис. 2.1.10. Переносное устройство с датчиком отпечатков пальцев

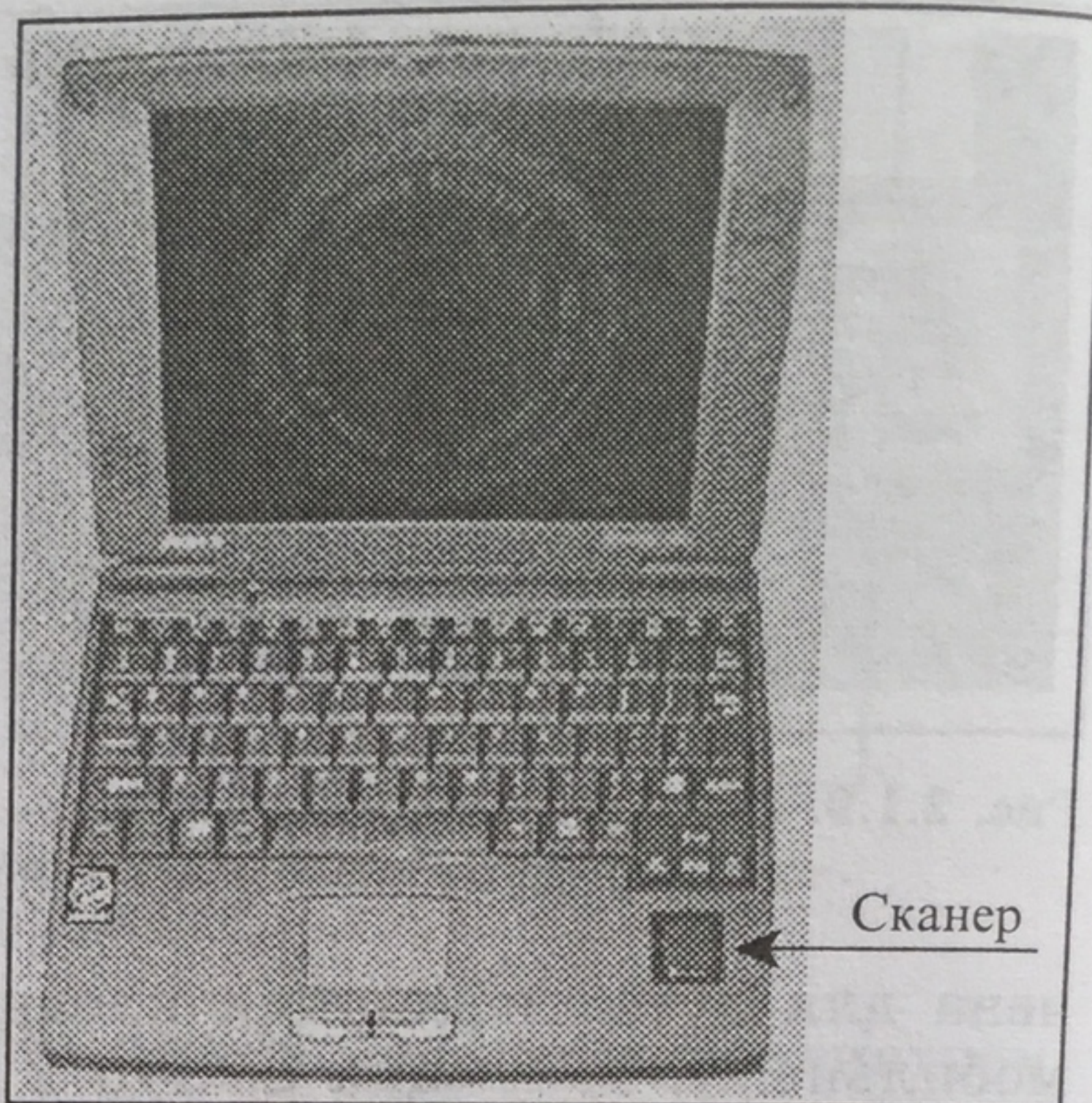


Рис. 2.1.11. Сканер отпечатков пальцев, встроенный в компьютер

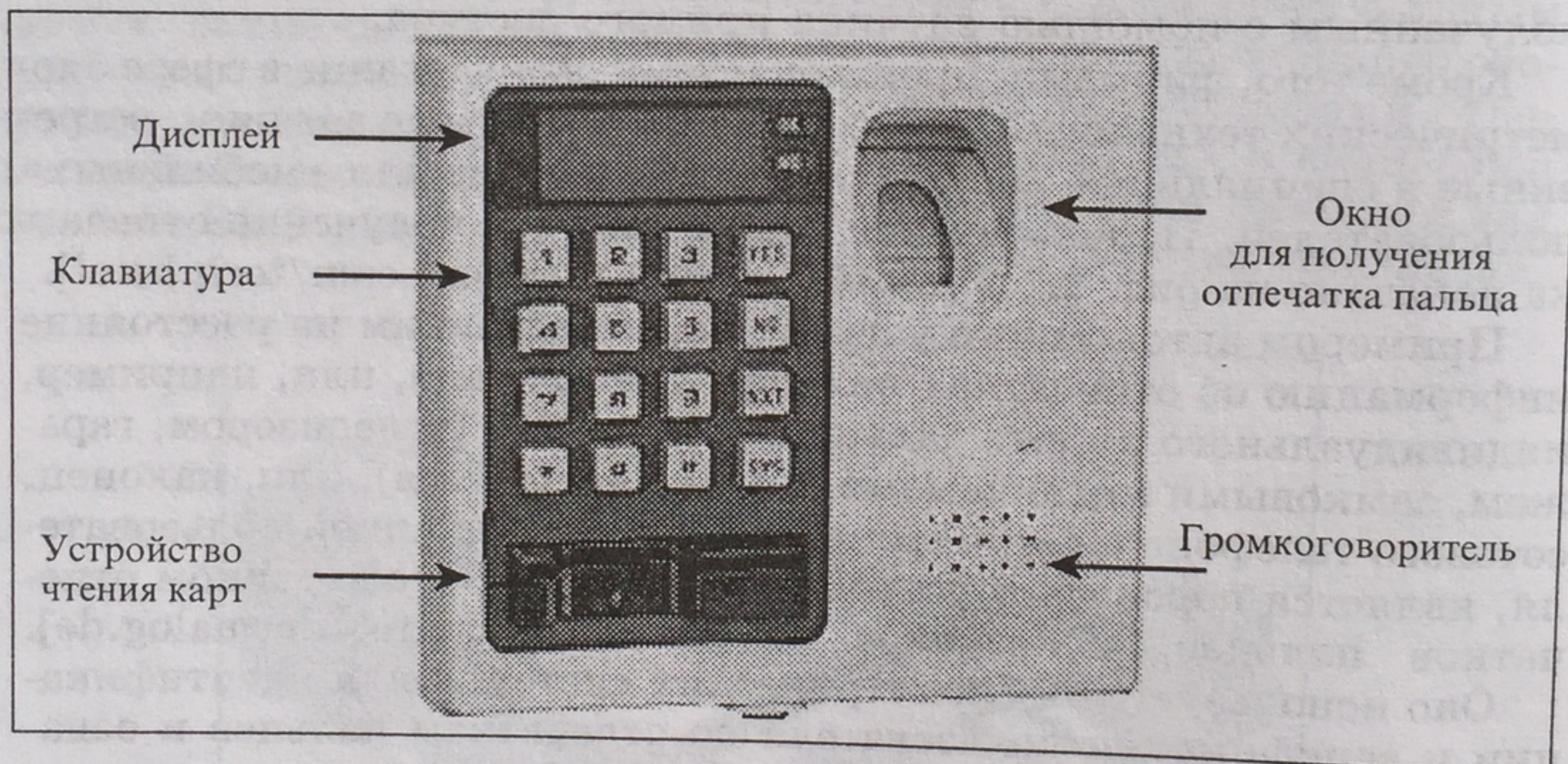


Рис. 2.1.12. Терминал верификации отпечатков пальцев FIT-100:

**Размеры:**

Терминал 148 × 135 × 60 мм  
Монитор 134 × 32 точек

**Регистрация:**

Время < 1,0 sec  
Память RAM 32 кбайт  
Емкость памяти до 3000 отпечатков

**Верификация:**

Время 0,3 sec  
FRR < 1,0 %  
FAR < 0,1 %

**Дополнительные возможности:**

управление запорными механизмами дверей;  
дополнительные порты — 3 для датчиков положения двери;  
клавиатура — 10 численных клавиш, 6 функциональных клавиш



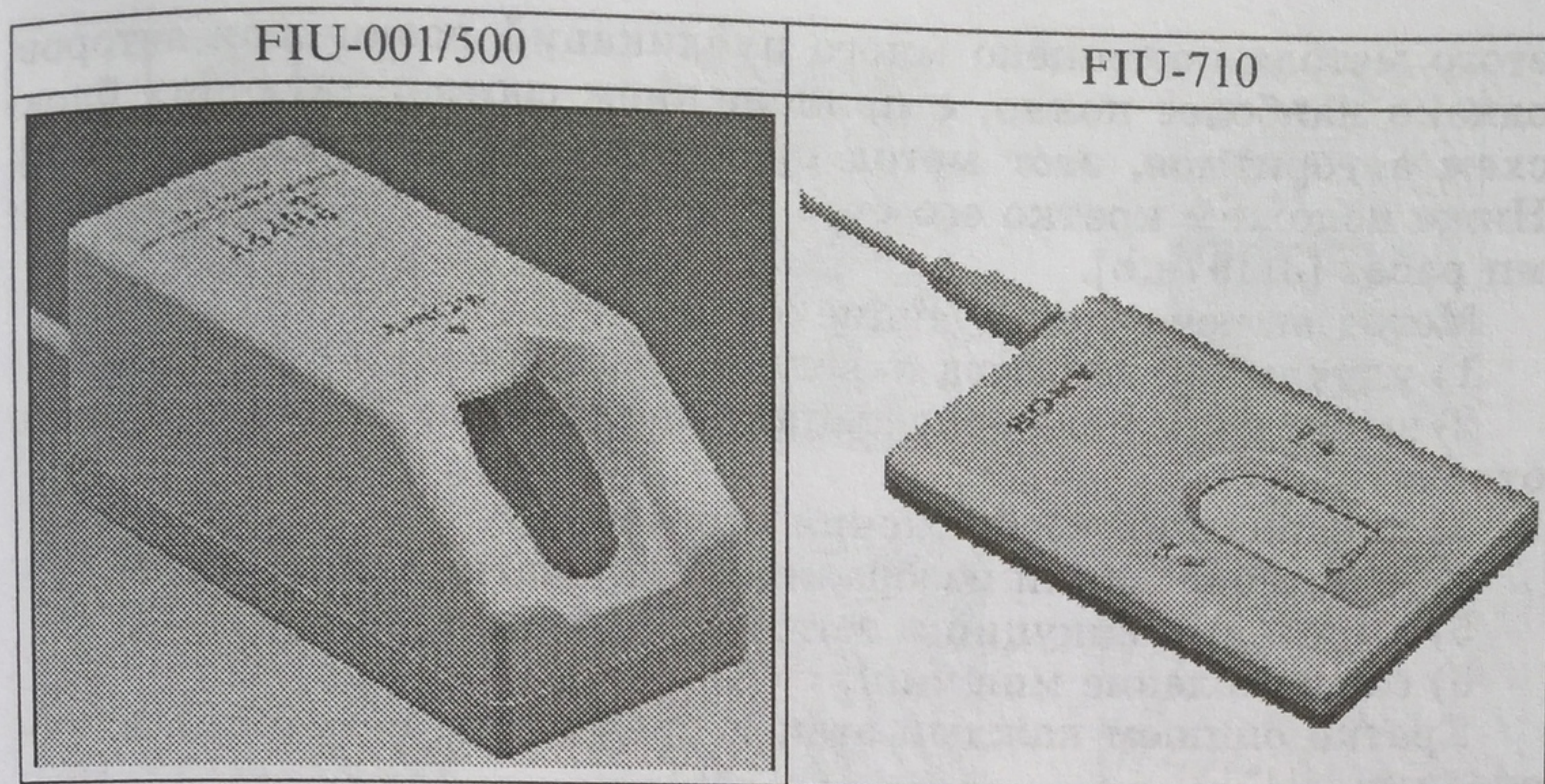


Рис. 2.1.13. Устройства верификации отпечатков пальцев для систем контроля доступа к помещениям

|                        | Устройство FIU-001/500       | Устройство FIU-710    |
|------------------------|------------------------------|-----------------------|
| Размеры                | 42 × 42 × 92,5 мм            | 9,5 × 54 × 85,6 мм    |
| Регистрация:           |                              |                       |
| Время                  | <1,0 с                       | 1,0 с                 |
| Память RAM             | 4 Мбайт                      | 512 кбайт             |
| Размер файла отпечатка | 256 байтов на отпечаток      | —                     |
| Эталон                 | —                            | 512 байт на отпечаток |
| Размер отпечатка       | —                            | 128 × 192 пикселей    |
| Емкость памяти         | 1000 отпечатков              | —                     |
| Верификация:           |                              |                       |
| Время                  | 0,3 с                        | 1 с                   |
| FRR                    | <1,0 % (по двум результатам) | 1 %                   |
| FAR                    | <0,1 %                       | 0,1 %                 |

(в зависимости от приложений). Память позволяет сохранять до 10000 эталонов. Независимость качества сканирования и получаемого результата от поворота пальца на сканере в плоскости XY. Способ сравнения с эталоном: один с одним или один с многими.

В заключение настоящего раздела скажем об отдельном классе устройств, включающих в себя как сканеры отпечатков пальцев, так и системы AFIS. Это терминалы и устройства верификации отпечатков пальцев, которые предназначены для систем контроля доступа к закрытым помещениям. С этой целью они содержат встроенные системы управления запорными механизмами дверей и порты для подключения датчиков положения двери. Внешний вид, например, терминалов фирмы Sony и их характеристики представлены на рис. 2.1.12 и 2.1.13.

#### 2.1.4. СРАВНЕНИЕ ОТПЕЧАТКОВ ПАЛЬЦЕВ

В технической литературе, посвященной проблемам автоматической идентификации и верификации отпечатков пальцев (в системах AFIS), наиболее часто упоминается метод, основанный на выделении и сопоставлении минучий. Описанию и применению



этого метода посвящено много публикаций различных авторов, однако наиболее полно, с приложением соответствующих блок-схем алгоритмов, этот метод представлен в работах [Jai97a,b]. Ниже изложим кратко его суть в соответствии с представлениями работ [Jai97a,b].

Метод включает следующие этапы:

1) улучшение качества исходного изображения отпечатка;  
2) вычисление поля ориентации папиллярных линий исходного отпечатка;

3) бинаризация изображения отпечатка;

4) утончение линий изображения отпечатка;

5) выделение минуций и запись их параметров в вектор;

6) сопоставление минуций.

Кратко опишем каждый этап и представим некоторые иллюстрации.

*Улучшение качества исходного изображения отпечатка.* Обычно с этой целью реализуется либо низкочастотная фильтрация, либо нормализация, призванные поправить структуру гребней папиллярных линий и улучшить резкость их границ.

*Вычисление поля ориентации папиллярных линий.* При этом все изображение разбивается на блоки, например,  $B \times B$  пикселей ( $B \geq 5$ , целое) и для каждого блока (по градиентам яркости в блоке) вычисляется один угол  $\theta$  ориентации папиллярных линий. Пример результата вычисления поля ориентации для фрагмента отпечатка пальца представлен на рис. 2.1.14.

*Бинаризация изображения отпечатка.* Процедура реализуется пороговой обработкой.

*Утончение линий изображения отпечатка.* Выполняется процедурой утончения линий бинарного образа до тех пор, пока эти линии не станут равными одному пикселю. Результат двух последних операций показан на рис. 2.1.15 на образах, отмеченных цифрами 2 и 3.

*Выделение минуций.* Реализуется локальной обработкой всего изображения отпечатка пальца с помощью маски  $9 \times 9$  пикселей и подсчетом числа пикселей, находящихся вокруг центра маски и имеющих ненулевые значения. Пиксель в центре маски принимается за минуцию, если он сам имеет ненулевое значение и если

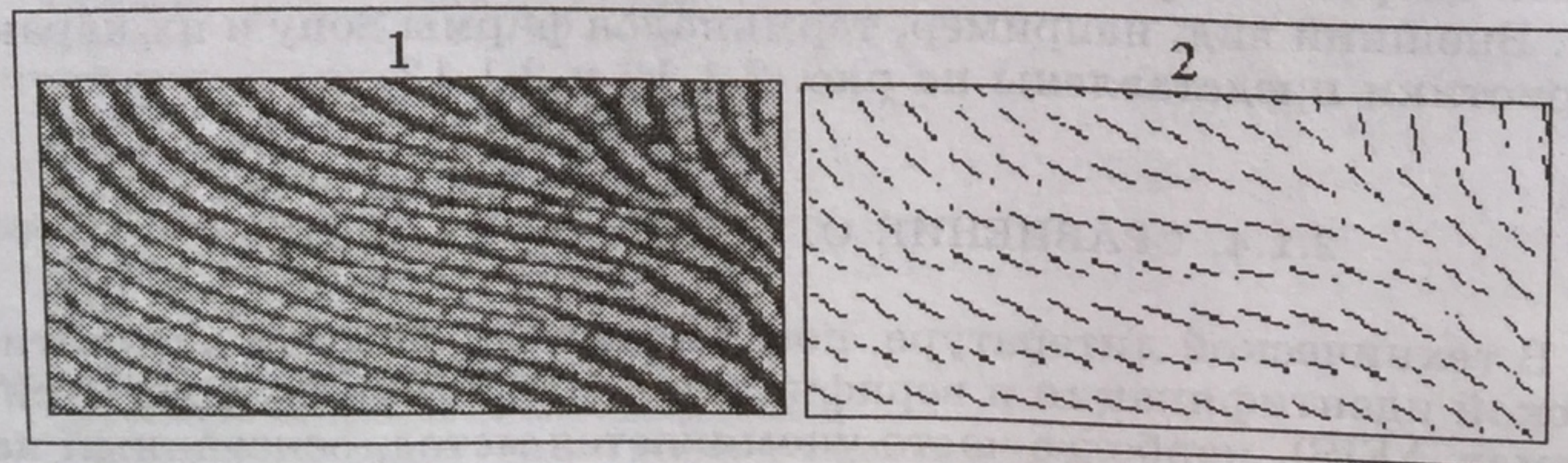


Рис. 2.1.14. Фрагмент отпечатка (1) и поле ориентации (2) его папиллярных линий



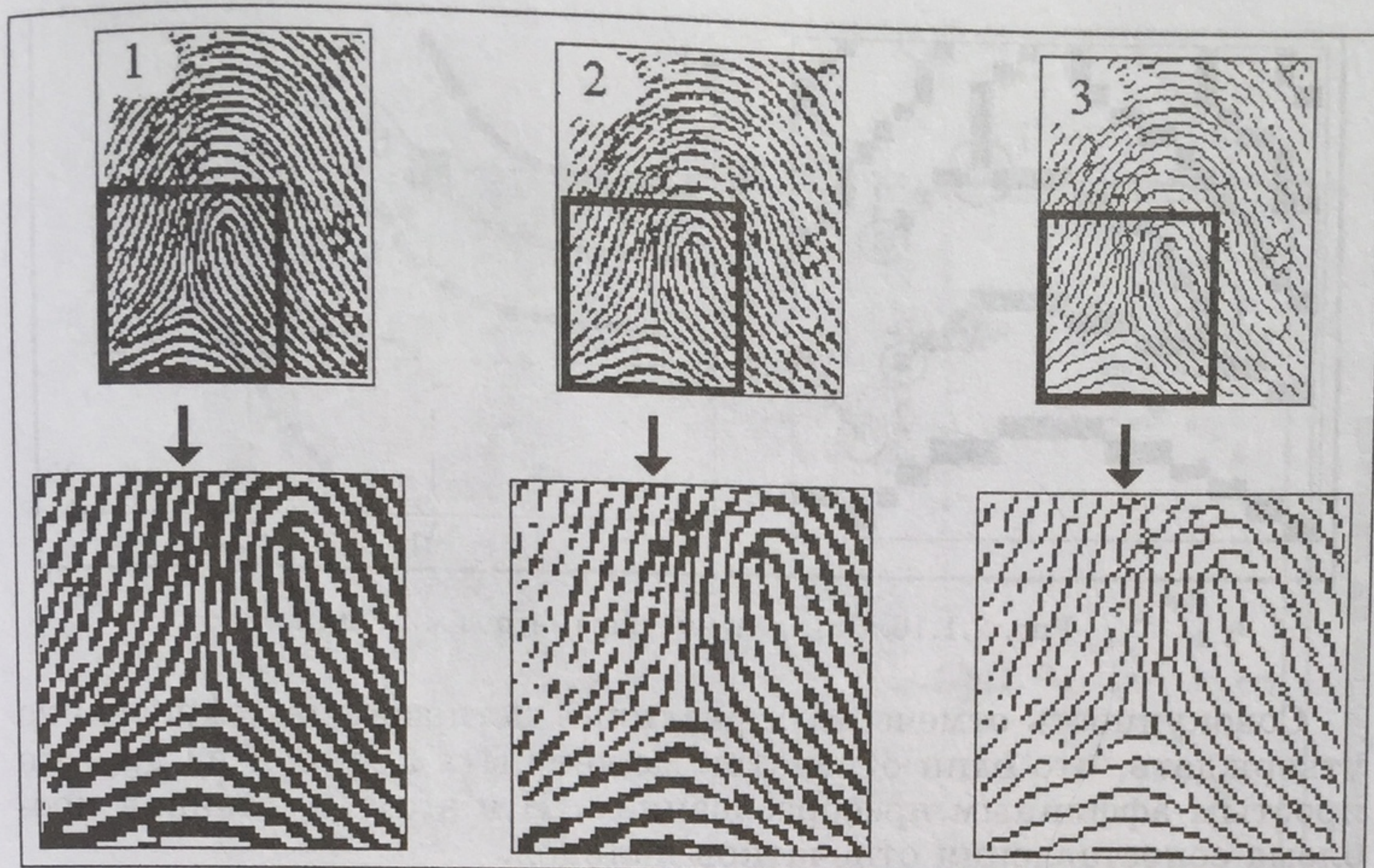


Рис. 2.1.15. Результаты бинаризации и утончения линий изображения отпечатка пальца

число «соседей» (например, величина  $S$ ) также ненулевое и равно 1 или 2. При этом, если  $S = 1$ , то обнаруженная минуция является минуцией типа «окончание» гребня, а при  $S = 2$  тип обнаруженной минуции — «раздвоение» гребня (рис. 2.1.16). Пример определения минуций, их типа и угла ориентации представлен на рис. 2.1.17. В рассматриваемом алгоритме этот угол определяется по вычисленному заранее полю ориентации в соответствии с координатами найденной минуции.

Координаты обнаруженных минуций, а также углы их ориентации записываются в вектор минуций в следующей форме:

$$W^{(p)} = \left[ (x_1, y_1, \theta_1), (x_2, y_2, \theta_2), (x_3, y_3, \theta_3), \dots, (x_p, y_p, \theta_p) \right], \quad (2.1)$$

где параметр  $p$  определяет число минуций.

На этапе регистрации пользователей системы AFIS вектор (2.1) определяет эталон  $E^{(p)}$ , который записывается в базу эталонов. На этапе распознавания вектор (2.1) определяет текущий отпечаток  $O^{(q)}$  и его параметры.

Заметим, что в реальной ситуации два отпечатка одного и того же пальца, полученные один за другим на одном и том же сканере, будут отличаться друг от друга поворотом относительно одной из осей, смещением узора отпечатка относительно осей  $X$  и  $Y$ , несколько различной деформацией узора (или нелинейными искажениями узора), возможными изменениями масштаба представления отпечатков и, наконец, разными периферийными участками реального пальца (рис. 2.1.18, верхняя часть).



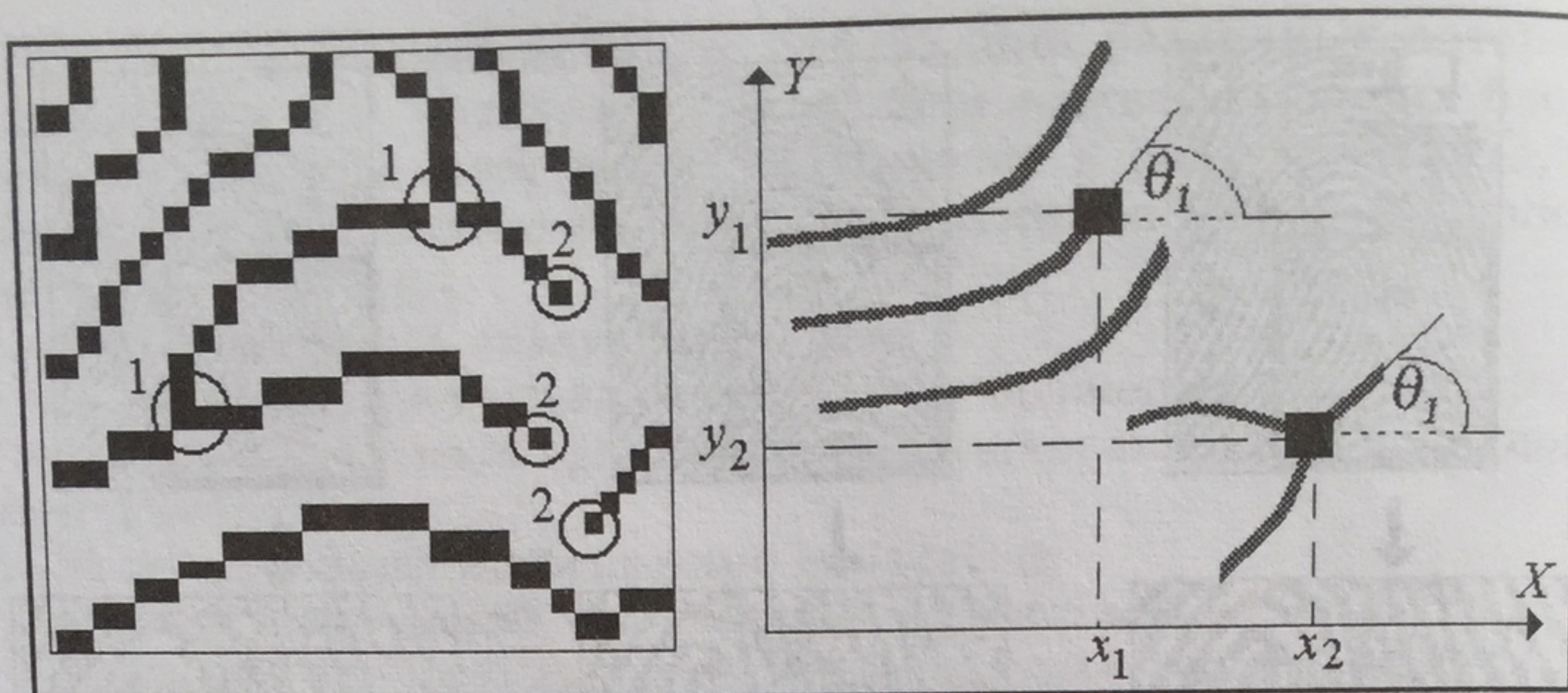


Рис. 2.1.16. Определение параметров минуций

Совокупность отмеченных различий не позволяет однозначно утверждать, что один отпечаток может быть получен из другого простым аффинным преобразованием. И в этом — главная проблема сопоставления отпечатков пальцев.

В приложении к записи отпечатков в форме (2.1) отмеченное выше различие в отпечатках означает, что  $p \neq q$  и, следовательно, векторы  $E^{(p)}$  и  $O^{(q)}$  имеют различную длину. Очевидно, что эти векторы содержат и некоторое число минуций, вообще не соответствующих друг другу. Изменения в пространстве минуций для пяти отпечатков одного и того же пальца отчетливо видны на рис. 2.1.18 (см. нижнюю часть рисунка).

В связи с отмеченным процесс сопоставления минуций должен быть реализован не за одну итерацию комплектно для всей их совокупности, а итерационно для каждой отдельной минуции. И в этом заключается главная проблема сопоставления минуций.

С учетом изложенного выше и в соответствии с методологией работ [Jai97a,b] процесс сопоставления минуций состоит из трех частей:

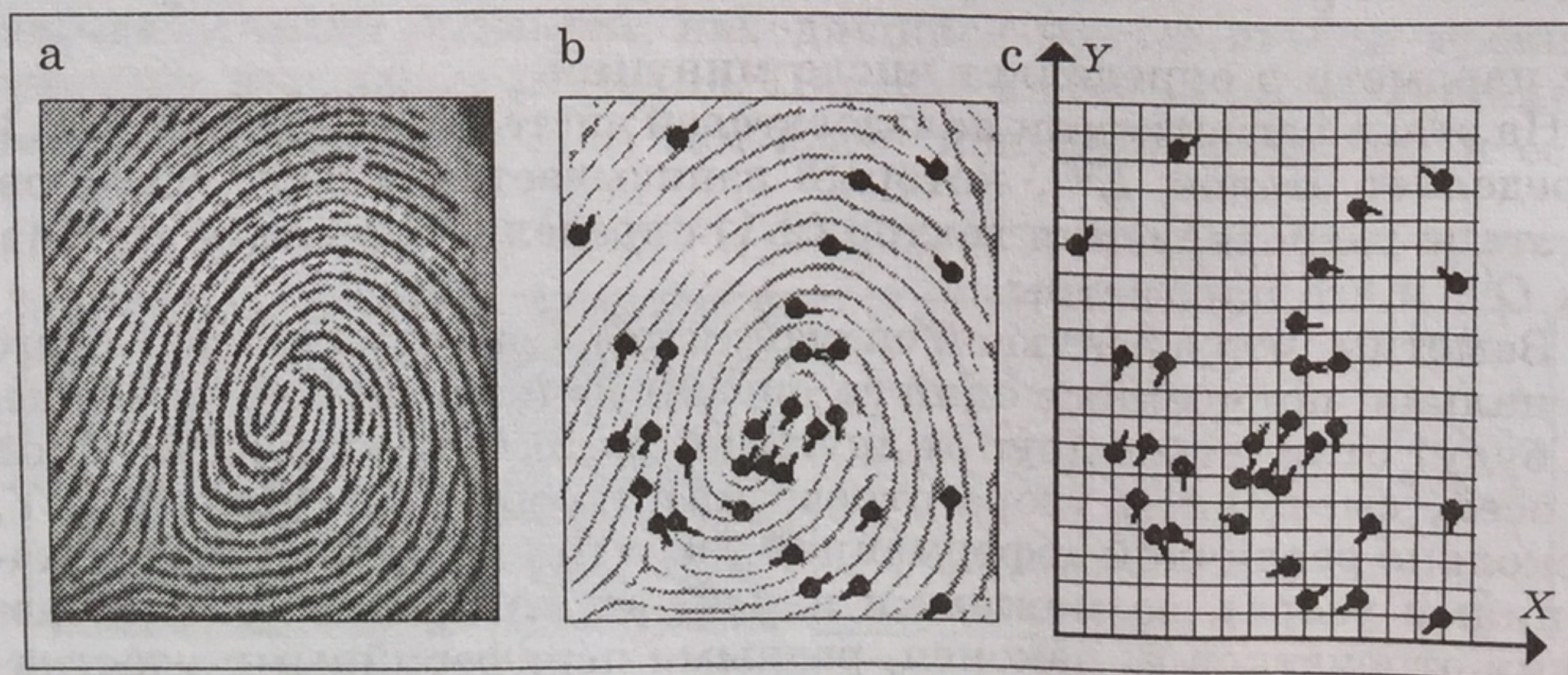


Рис. 2.1.17. Определение минуций (а), их типов (b) и углов (с) ориентации



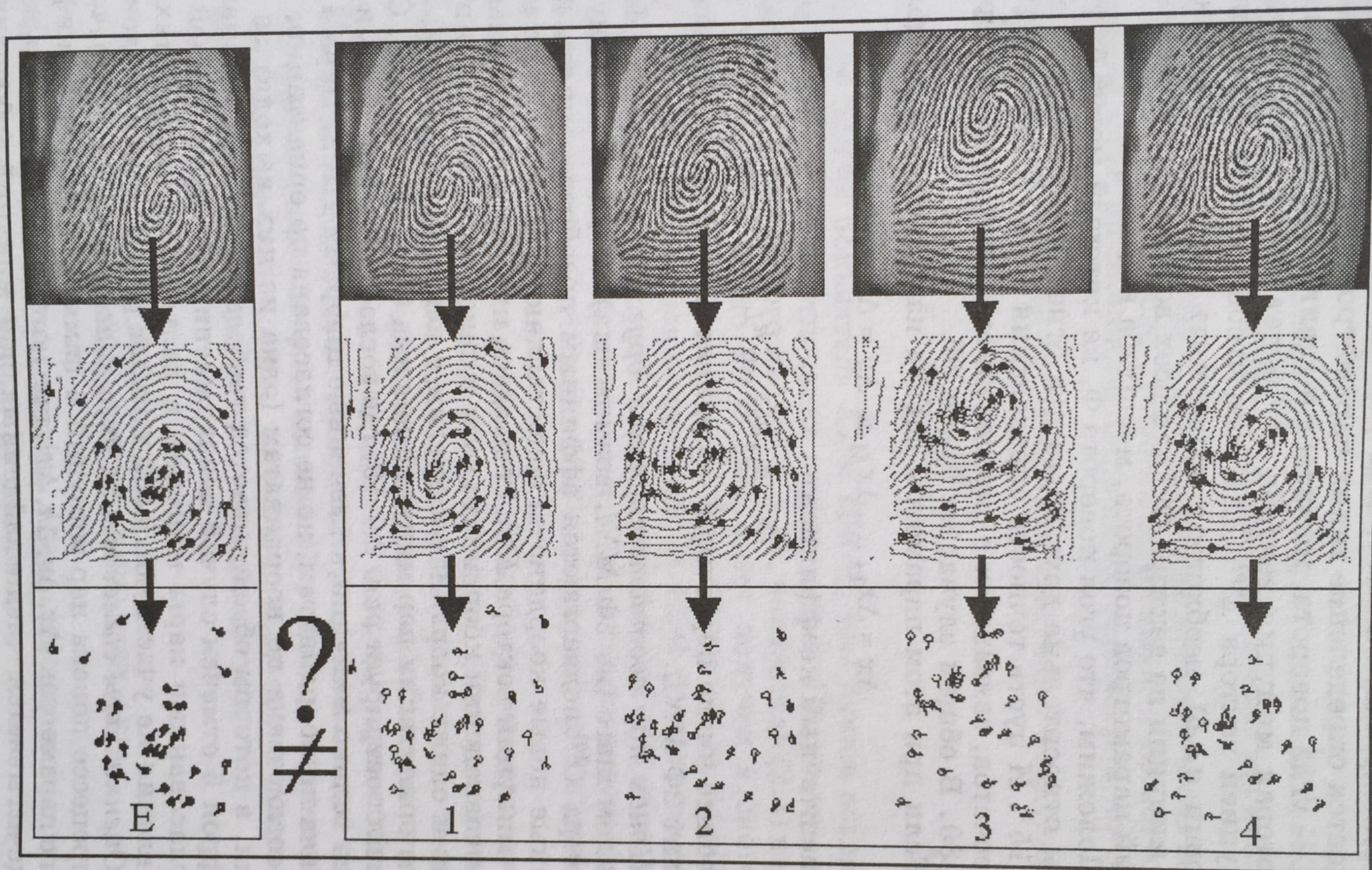


Рис. 2.1.18. Отпечатки одного и того же пальца и соответствующие им минущии



- 1) регистрация данных;
- 2) поиск пар соответствующих друг другу минуций;
- 3) оценка меры согласования двух сопоставляемых отпечатков.

*Регистрация данных.* Под регистрацией в данном случае понимается определение таких параметров аффинных преобразований — угла поворота, масштаба и сдвигов ( $\Delta x$  и  $\Delta y$ ), при которых некоторая минуция вектора  $O^{(q)}$  будет согласована с некоторой минуцией вектора  $E^{(p)}$ . Для того чтобы решить эту задачу (при условии  $q \neq p$ ), необходимо выполнить пересчет координат каждой минуции из вектора  $O^{(q)}$  для всех возможных изменений значений параметров поворота, масштаба и сдвигов.

Положим, что угол поворота  $\Phi$ , на который могут быть повернуты относительно друг друга два отпечатка пальца, составляет  $\Phi = 15^\circ$ . И пусть точность определения угла составляет  $1^\circ$ . Тогда пересчетов, связанных с определением только угла поворота, будет 30. В общем случае их будет  $2\Phi$ .

Если при этом параметры сдвига изменяются в пределах

$$\Delta x = \Delta x_1, \dots, \Delta x_M \text{ и } \Delta y = \Delta y_1, \dots, \Delta y_N, \quad (2.2)$$

а масштабный коэффициент

$$k = k_1, k_2, \dots, k_L, \quad (2.3)$$

то общее число пересчетов координат одной только минуции составит  $2\Phi MNL$ .

*Поиск пар соответствующих друг другу минуций.* Итак, на каждом шаге (из  $2\Phi MNL$  возможных) координаты минуции из вектора  $O^{(q)}$  подвергаются аффинным преобразованиям, и полученные новые координаты сопоставляются с каждой из координат минуций вектора  $E^{(p)}$ . Речь здесь идет не о пунктуально точном совпадении координат (хотя и такое возможно), а о приближении одних координат к другим с некоторой ошибкой, но при этом ориентация пар минуций должна быть одинаковой. Сказанное иллюстрирует рис. 2.1.19, где согласованными парами считаются только минуции, отмеченные цифрами 1 и 2. Пара 3 согласована по координатам, но не согласована по ориентации; пара 4 не согласована по координатам (одна из них выходит за рамки поля, в котором обеспечивается требуемая точность) и, наконец, цифрой 5 отмечен случай ненахождения соответствующей пары. Согласованные пары минуций в дальнейших итерациях сопоставления не участвуют (если не идет речь о новом эталоне).

*Оценка меры согласования двух сопоставляемых отпечатков.* В процессе поиска пар минуций в памяти сохраняется информация о параметрах  $\{\Delta x, \Delta y, \Delta k, \Delta \Phi\}$ , при которых найдена пара, а также подсчитывается число найденных пар минуций (обозначим это число буквой  $D$ ).

Если одинаковая информация о параметрах  $\{\Delta x, \Delta y, \Delta k, \Delta \Phi\}$  повторяется довольно часто, то это свидетельствует о высоком качестве согласования отпечатков и определении истинных парамет-



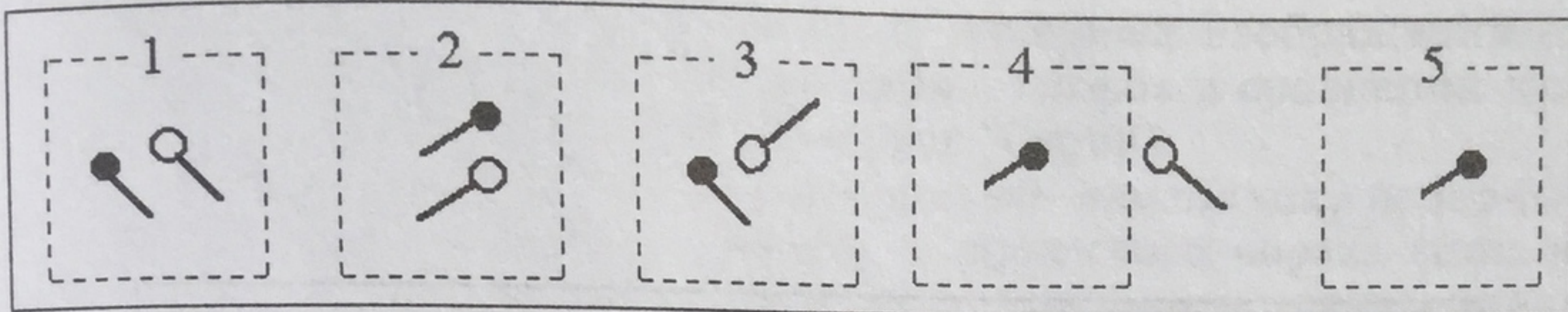


Рис. 2.1.19. Варианты решений поиска пар при сопоставлении минуций

ров аффинных преобразований, связывающих два отпечатка между собой. В этом случае число итераций согласования может быть уменьшено за счет применения найденных параметров ко всем оставшимся минуциям.

Количественную оценку согласования получают следующим образом [Jai97]:

$$K_O = \frac{D \cdot D \cdot 100\%}{p \cdot q} \quad (2.4)$$

Ясно, чем ближе показатель  $K_O$  к 100 %, тем лучше согласуются два сравниваемых отпечатка между собой. Например, для результата, представленного на рис. 2.1.18, показатель (2.4) находится в пределах от 65 до 80 % (результат получен в системе «VeriFinger v3.3» для отпечатков из базы Biometric System Lab [Fin00]). Наивысшее значение показателя (2.4) соответствовало сравнению эталона (на рисунке отмеченного буквой E) с образами 1 и 2, а наименьшие — с образами 3 и 4. Как показали результаты, выдаваемые системой «VeriFinger v3.3» во всех четырех случаях, верификация прошла успешно.

Если согласование выполнялось в задаче верификации отпечатков, то на этом весь процесс заканчивается. Если же согласование выполнялось в задаче идентификации отпечатков, то описанный процесс надо повторять для всех эталонов базы минуций. В некоторых системах идентификации из базы выбирают 10 эталонов, которым соответствуют оценки (2.4), превышающие некоторый заданный порог. За окончательный результат принимается тот вариант идентификации, для эталона которого получено максимальное значение оценки (2.4) из этих десяти.

#### 2.1.5. НЕСКОЛЬКО СЛОВ О ДРУГИХ ПОДХОДАХ К РЕАЛИЗАЦИИ AFIS

Несмотря на высокую надежность представленного выше метода распознавания отпечатков, до настоящего времени не прекращаются поиски еще более совершенных подходов к реализации AFIS. Представим еще два метода, реализующие сравнение двух отпечатков пальцев между собой:

- основанный на использовании глобальных признаков [Rat96];



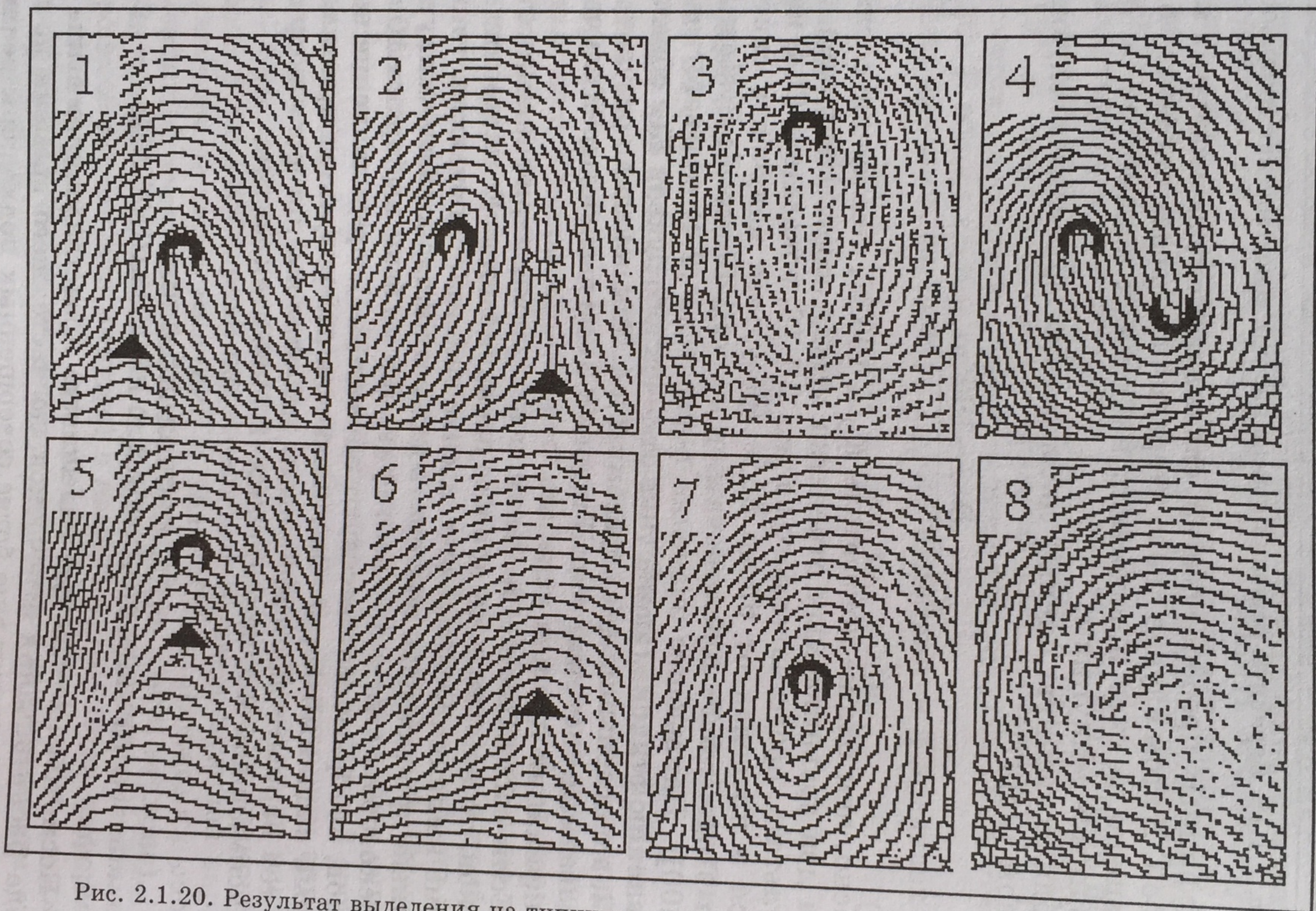


Рис. 2.1.20. Результат выделения на типичных узорах признаков «ядро» и «дельта» (дуга):  
 ◐ — ядро; ▲ — дельта (дуга)

пе  
ан

яв.

па

та»

ляе

при

нун

печ

пал

стал

соот

под

чат

осно

печ

ми

сам

цев

1

6

Рис. 2.1



• основанный на преобразованиях исходных изображений отпечатков пальцев в форму специальных графов и сравнении их с аналогичными графами базы эталонов [Сар99].

В методе, основанном на глобальных признаках, основным является обнаружение на области выделенного образа (или на папиллярных узорах) таких признаков, как ядра и пункты «дельта». Число таких признаков и их взаимное расположение позволяет однозначно классифицировать тип узора. Расстояние между признаками, наклон соединяющих их линий дают дополнительную информацию для классификации каждого конкретного отпечатка.

Пример выделения признаков на типичных узорах отпечатков пальцев представлен на рис. 2.1.20 [Rat96]. Классификация представленных папиллярных узоров обсуждалась в разделе 2.1.1 и соответствует узорам на рис. 2.1.1. Представленный в [Rat96] подход использован для предварительной классификации отпечатков пальцев. Окончательное распознавание реализовано на основе локальных признаков. Такой итерационный подход обеспечивает режим реального времени при работе с большими базами данных.

Определение типа узоров на пальцах может представлять и самостоятельный интерес, поскольку по папиллярным узорам пальцев можно узнать характер, темперамент и способности человека,

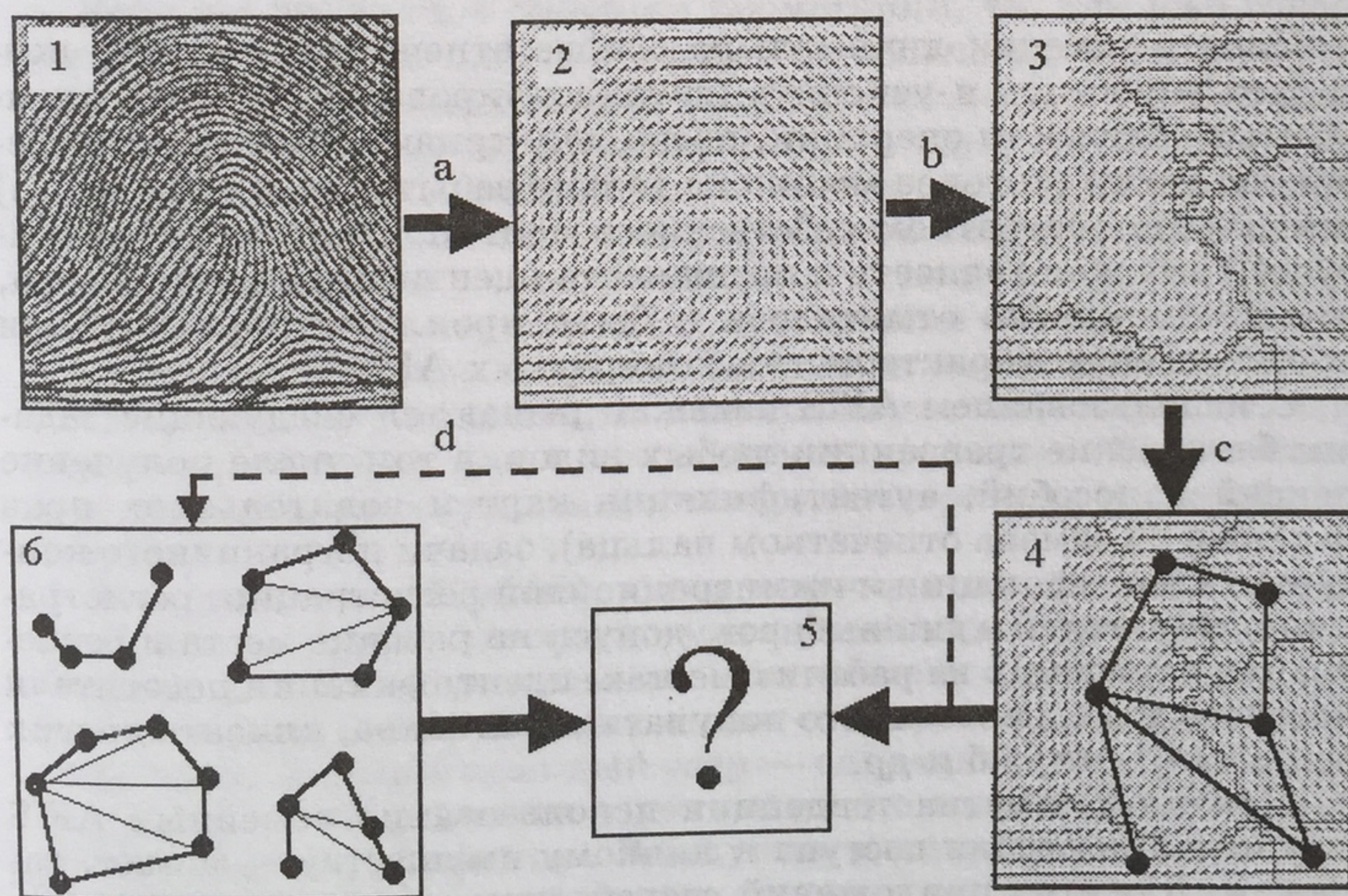


Рис. 2.1.21. Преобразование изображения отпечатка пальца в форму специального графа и определение меры его подобия с эталонами



а также определить, что ждет человека в будущем [Пер98]. В решении и этих задач может быть полезен подход, представленный в [Rat96].

Идеи второго метода, основанного на «графовом представлении» отпечатков пальцев, иллюстрирует схема, показанная на рис. 2.1.21 [Cap99].

В соответствии с этой схемой каждое исходное изображение 1 отпечатка пальца преобразуется в новое изображение 2, представляющее поле ориентации папиллярных линий. На таком изображении легко заметить области с неизменной ориентацией папиллярных линий. Изображение «с ориентацией» делится на эти области по своим границам, в результате чего получается изображение 3. Далее определяют центры выделенных областей, соединяя которые получают некоторый уникальный граф (см. изображение 4), представляющий конкретный отпечаток пальца. Описанный процесс отмечен стрелками «а» — «с». Стрелкой «d» отмечена запись графа в базу эталонов (блок 6) на этапе регистрации отпечатков или в базу текущих данных в процессе распознавания нового отпечатка.

Определение меры подобия вновь полученных графов с графами из базы эталонов реализуется отдельным блоком 5.

### Вместо заключения

Идентификация личности на основе отпечатков пальцев позволяет упростить и ускорить (автоматизировать) многие выполняемые человеком операции, поскольку тривиальное удостоверение личности (которое можно не только забыть, но и утратить...) заменяется отпечатком собственного пальца. При этом гарантируется высокая точность и надежность идентификации человека, что неоднократно отмечалось и было проиллюстрировано при представлении характеристик конкретных AFIS.

С использованием AFIS надежно решаются следующие задачи: банковские транзакции любых видов, в том числе получение пенсий и пособий, аутентификация карт и водительских прав (а также их замена отпечатком пальца), задачи пограничного контроля/идентификации и иммиграционной регистрации, регистрация избирателей в дни выборов, допуск на рабочие места и регистрация служащих на рабочих местах, идентификация посетителя частного дома, постоянного покупателя магазина, клиента других социальных служб и др.

При этом заметна тенденция использования семейных AFIS для целей контроля доступа к личному имуществу — домам, гаражам. Для этих приложений разработаны автономные компактные и в то же время функционально универсальные терминалы идентификации (аутентификации) человека по отпечаткам паль-



ца. Такие терминалы включают кроме основного устройства — сканера отпечатка пальца — много дополнительных устройств, и в том числе системы управления замковыми механизмами дверей, интерфейсы для подключения датчиков положения дверей, считыватели карт, клавиатуру для ввода паролей и PIN'ов, что существенно облегчает их установку и обслуживание, а также гарантирует широкое применение. Вот уж действительно, наступает эра, когда палец становится ключом [Мат00]. И сегодня уже говорят о защите доступа к сугубо личным предметам (опасным в чужих руках), таким, например, как оружие.

Криминалистика и правосудие — также известные сферы приложения AFIS, что отмечалось уже выше.

Интенсивно и экстенсивно расширяется применение биометрических систем класса AFIS и в сфере высоких технологий, где решаются две основные задачи — контроль доступа к информационным ресурсам и идентификация их пользователей. Причем появились так называемые «мобильные пользователи», реализующие доступ к информационным ресурсам по радиотелефонным линиям. Их идентификация выполняется как бы «на ходу», например через системы сотовых телефонов. При этом заметна и тенденция защиты персональных информационных ресурсов — персональных компьютеров (в том числе ноутбуков) и доступа к ним, а также доступа к ресурсам через мобильные каналы, например через сотовые телефоны. С использованием мобильных AFIS эти задачи легко решаются.

Если же говорить о будущем биометрики, то, как уже говорилось, по узорам на пальцах можно не только идентифицировать человека, но и предсказывать, будущее человека. «Ученых давно интересовали узоры на пальцах, которые характеризуются своей уникальностью и неизменностью в течение жизни. Наука, которая занимается считыванием и интерпретацией информации с пальцевых (папиллярных) узоров, называется дерматоглифика» [Пер98].

Если говорить о перспективах применения технологии отпечатков пальцев, то очень скоро она выйдет за рамки простых систем AFIS, «поскольку на кончиках пальцев можно читать так же, как в медицинской карте или в милицейском протоколе. По папиллярным узорам можно узнать характер, темперамент и способности человека».

Как отмечено в [Пер98], уже в течение нескольких лет дерматоглифику используют при изучении возможностей спортсменов. Сегодня уже точно установлено, что человек с узором «дуга» добьется победы только в таких видах спорта, как коньки-спринт и шорт-трек. А самый сложный узор — спираль (завитки) или двойные петли — встречаются у чемпионов по вольной борьбе, тяжелой атлетике, фехтованию. В будущем ученые надеются расшифровать назначение и детали всех узоров.

Для реализации подобных проектов имеется буквально все — датчики, устройства и терминалы для получения отпечатков паль-



цев, методы, алгоритмы и программное обеспечение для выделения и классификации папиллярных узоров, их деталей и базы данных, представляющие собой материал для анализа, в том числе папиллярные узоры известных исторических и популярных личностей и т. д.

## 2.2. ТЕХНОЛОГИЯ, ИСПОЛЬЗУЮЩАЯ ГЕОМЕТРИЮ РУКИ

«Рука — вышедший наружу мозг», — писал философ Кант. Что он хотел этим сказать? А именно то, что руки отражают наш внутренний мир. При этом ладонь — инструмент, который у нас «всегда под рукой» в прямом и переносном смысле [Лев99]. И практически все о конкретном человеке можно прочесть по его руке.

Однако в биометрике в целях идентификации (или верификации) человека используется пока только простая геометрия руки — ее размеры и форма, а также некоторые информационные знаки на тыльной стороне руки. К последним относятся образы на сгибах между фалангами пальцев и узоры расположения подкожных кровеносных сосудов [Gol97, San00].

Как следует из отчета международной группы по биометрике (BIG), технология идентификации по геометрии руки использовалась в первых коммерческих биометрических системах фирмы Identimat (1976 г.). Сегодня две основные компании предлагают на рынке свои системы этого класса: BioMetPartners и Recognition Systems Inc ([www.recogsys.com](http://www.recogsys.com)). Первая компания производит систему верификации «Digi-2», в которой используется информация о размерах и форме двух пальцев руки. Вторая компания производит систему «ID3D HandKey», в которой используется трехмерный образ четырех пальцев и части руки. Последняя система относится к системам третьего поколения. Обе компании используют свои разработки в системах контроля доступа.

Вообще же с руки можно «снять» до 90 информационных знаков. В биометрике, например, совершенно не используется информация об узоре на ладони, а она также уникальна (рис. 2.2.1, левая часть), как уникальна судьба и жизнь каждого человека. Возможно, за этим — будущее биометрики, использующей информационные знаки руки.

**Информационные знаки руки.** В биометрике, использующей геометрию руки, прослеживаются два основных подхода. Первый (существует уже более 25 лет — от зарождения биометрических систем контроля доступа к помещениям) основан исключительно на геометрических характеристиках кисти руки, второй (более современный) — на смешанных характеристиках геометрических и образных. К последним, как уже отмечалось, относятся образы на сгибах между фалангами пальцев, узоры (расположение) подкожных кровеносных сосудов.



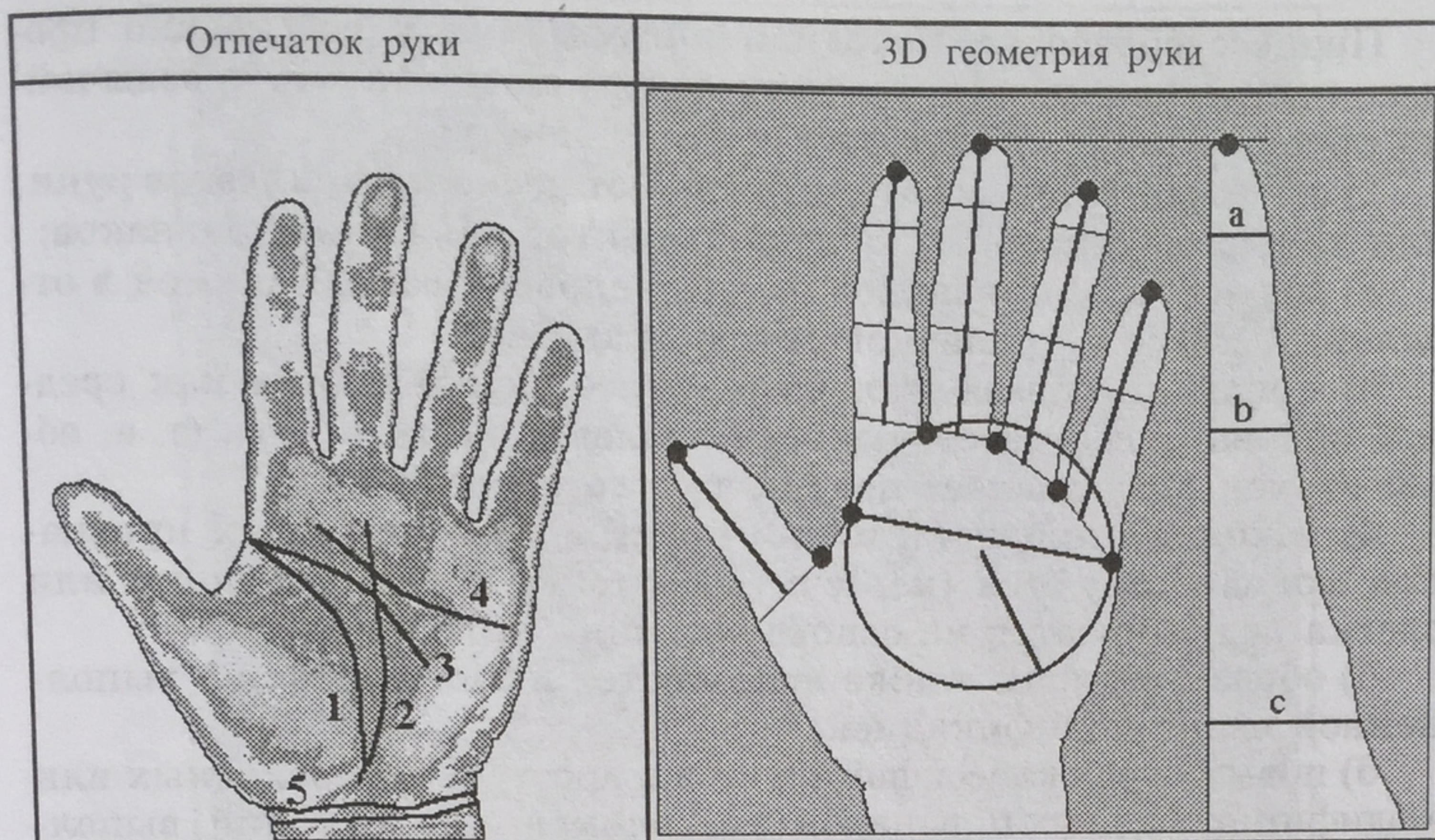


Рис. 2.2.1. Узор на ладони, состоящий из пяти важнейших линий (слева), контрольные точки и 17 геометрических признаков руки (справа)

Суть первого подхода иллюстрирует рис. 2.2.1, где с левой стороны показан отпечаток руки, а с правой — 3D геометрия руки. В рассматриваемом подходе вся интересующая нас информация собрана в отличие от отпечатка руки в горизонтальном и вертикальном силуэте ее кисти.

На рис. 2.2.1 показаны контрольные (характеристические) точки силуэта руки и 17 исходных геометрических признаков руки [Gol97], в данном случае отмеченные отрезками прямых линий, которые не входят в силуэты кисти. Как видно, исходными биометрическими признаками руки являются ширина ладони, радиус вписанной в ладонь окружности, длины пальцев (определяемые как расстояния от выделенных верхних контрольных точек до середин линий, соединяющих нижние контрольные точки), ширина пальцев и высота кисти руки в трех пунктах, отмеченных линиями *a*, *b* и *c*.

В некоторых других подходах, в том числе в рассматриваемом в [San00], приведен 21 исходный геометрический признак руки, из которых 17 совпадают с признаками, показанными на рис. 2.2.1.

В системах верификации личности человека по геометрии руки используются не все, а только некоторые из этих исходных признаков, а также другие, полученные из исходных при их математической обработке [San00]. Такими признаками, например, могут служить углы между контрольными точками, средние значения и дисперсия значений исходных признаков.

Особенностью рассмотренного способа получения биометрических характеристик руки является его простота, а также то, что каждый эталон руки может быть записан очень компактно в форме некоторого вектора значений признаков.



Процесс обработки исходной информации и собственно процесс идентификации (аутентификации) людей также чрезвычайно прост [San00]. Он сводится к следующему:

- 1) от каждого пользователя получают несколько силуэтов руки, для каждого из которых получают вектор значений признаков;
- 2) все векторы признаков одного человека объединяются в отдельный класс в признаковом пространстве;
- 3) признаки образа-эталона в классе определяются как средние значения соответствующих признаков всего класса (т. е. образ-эталон представляет центр каждого класса);
- 4) исходные признаки и/или пространство исходных признаков модифицируются (пересчитываются в новые признаки или слегка редуцируются на основе селекции признаков);
- 5) образы-эталоны также изменяются в соответствии с выполненной выше модификацией;
- 6) новый образ классифицируется в пространстве исходных или модифицированных признаков (в зависимости от операций, выполненных в пунктах 4 и 5) при его сравнении с образом-эталоном.

Мерой подобия сравниваемых образов может служить, например, расстояние в метрике  $L_2$  между новым образом и центрами классов: чем меньше это расстояние, тем ближе новый образ к соответствующему классу.

Если в биометрической системе выполнялась процедура аутентификации, то оценивается расстояние между новым образом и известным эталоном, и при этом вычисленное расстояние не может превосходить заданное значение. В противном случае считается, что аутентификация не подтверждена.

Однако к применению подобных биометрических систем нужно относиться крайне осторожно, поскольку не составляет большого труда изготовить простой (например, из картона и пластика) муляж кисти руки с характеристиками санкционированного (зарегистрированного пользователя) и использовать этот муляж для несанкционированного доступа в систему.

Суть второго подхода — «снятие» с руки четырех характеристик, из которых три являются скалярами и относятся к размерам пальцев, а четвертая — представляет собой полутонное изображение складок кожи на сгибе между фалангами.

Способ получения всех четырех характеристик показан на рис. 2.2.2. Три первые характеристики — это ширина указательного пальца 1, высота указательного пальца 2 и длина среднего пальца 3, оцениваемая так, как показано на рис. Характеристика 4 в рассматриваемом случае представляет собой изображение складок кожи на сгибе между средней и нижней фалангами указательного пальца. Как утверждают некоторые рекламные источники компании Recognition Systems Inc, вся информация о руке в рассматриваемом классе систем может быть записана не более чем 9 байтами.



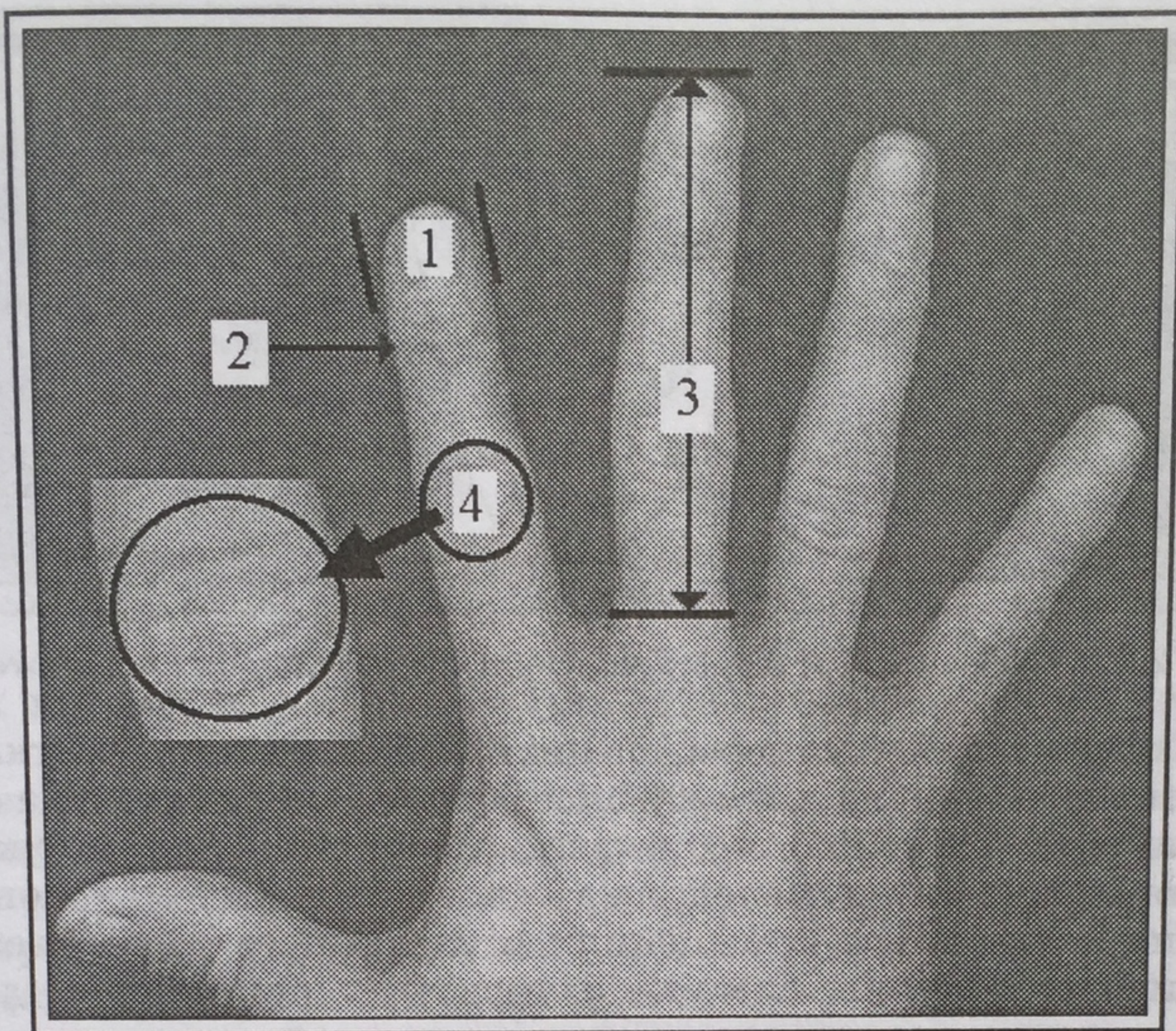


Рис. 2.2.2. Один из вариантов биометрических характеристик руки

Наличие биометрической характеристики руки в форме полутонового изображения существенно затрудняет изготовление муляжа, который может быть использован для взлома системы идентификации. При этом следует заметить, что фирмы, поставляющие на рынок коммерческие системы идентификации человека по его руке, не предоставляют конкретной информации о признаках (характеристиках) руки, используемых в своих системах.

Поэтому вместо варианта четвертой характеристики, показанной на рис. 2.2.2, в конкретной биометрической системе может быть выбрана какая-либо другая или несколько. Незнание этой информации также усложняет изготовление муляжа руки и уменьшает вероятность физического доступа к системе.

**Устройства для получения информации о руке.** Для получения параметров кисти руки существуют различные сканеры (рис. 2.2.3 позиции 1 и 3).

На рис. 2.2.3, позиция 2 показан способ фиксации руки в устройствах в процессе ее сканирования. И это, пожалуй, единственное специфическое требование при сканировании руки. Никаких других особых требований, в том числе к температуре, влажности и чистоте рук устройства этого класса не предъявляют ([www.recogsys.com/products/accessories](http://www.recogsys.com/products/accessories)). Последнее обстоятельство позволяет применять устройства рассмотренного класса в различных производственных условиях.

**Технология, использования геометрии руки.** Некоторая громоздкость представленных устройств связана с использованием



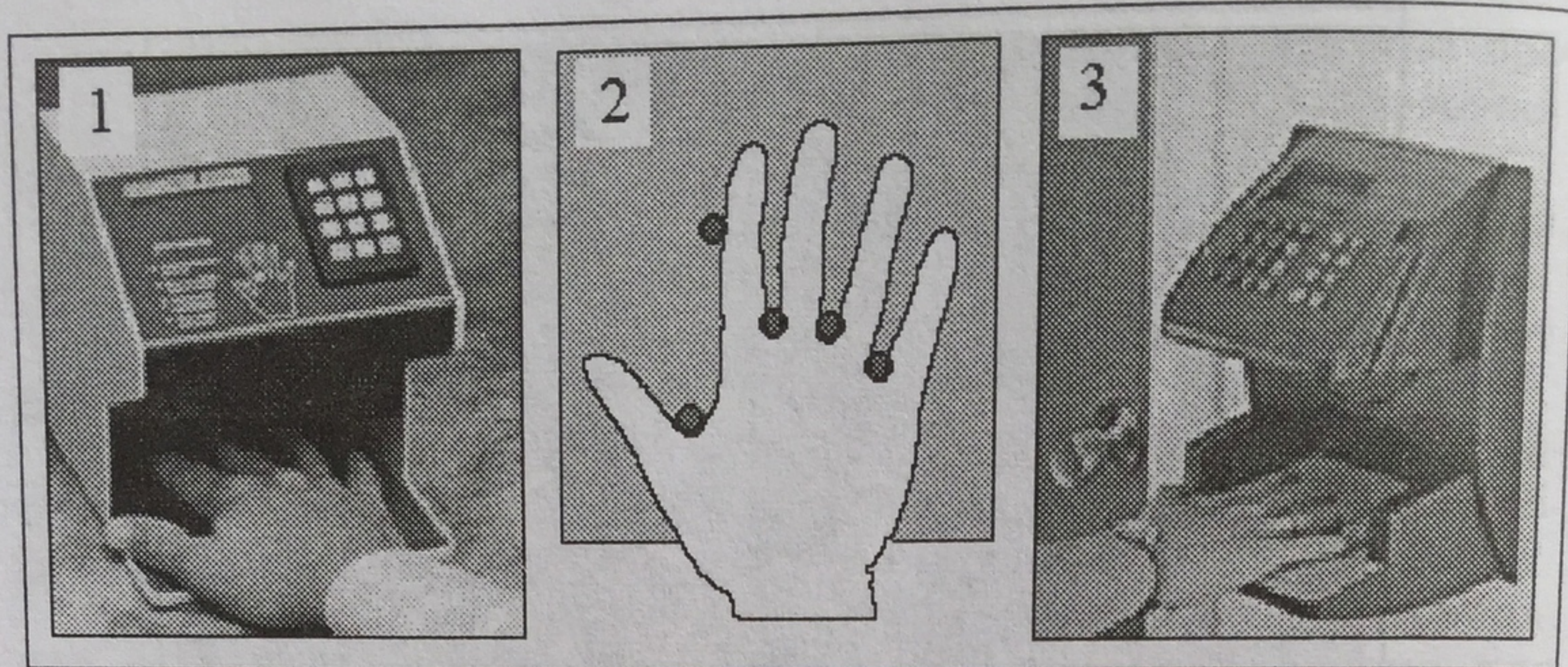


Рис. 2.2.3. Сканеры (1, 3) геометрии руки и способ фиксации в них руки (2)

в них телевизионной камеры, специальной системы подсветки руки и уголкового зеркала, отражающего боковой образ руки на объектив камеры. Возможно, что именно габариты сканеров рук определили сферу их применения (в основном в системах контроля доступа к закрытым помещениям и различного рода объектам).

Широко известны следующие примеры применения данной технологии на практике:

1) контролировался доступ в Олимпийскую деревню (г. Атланта, США) на Играх 1996 г.;

2) контролируется доступ в помещения и на объекты аэропортов, перемещения персонала по территории аэропортов, при получении багажа, контролируются пассажиры при выходах на посадку в самолеты. Например, в международном аэропорту в Сан-Франциско установлено более 90 сканеров геометрии рук;

3) работают системы идентификации (верификации) клиентов банков, а также системы контроля доступа в специальные помещения/хранилища в банках. Более 100 сканеров геометрии рук установлено в банках России;

4) контролируется доступ на объекты, а также перемещения персонала на атомных электростанциях (США, Япония, Канада);

5) функционируют системы аутентификации людей в службе иммиграции и натурализации США, используемые в рамках программы INSPASS.

### 2.3. ТЕХНОЛОГИЯ, ИСПОЛЬЗУЮЩАЯ РАСПОЗНАВАНИЕ ЛИЦ

Во введении отмечалось, что для интеллектуальных сред следующего поколения лучше всего подходят технологии идентификации человека по его лицу. Как отмечено в [Pen00], эти технологии «ненавязчивы (распознавание происходит на расстоянии, не задерживая и не отвлекая человека), они, как правило, пассивны



(не требуют специальных действий со стороны человека), не ограничивают пользователя в свободе перемещений и к тому же относительно недороги. Но самое важное, по-видимому, то, что люди обычно узнают друг друга по лицам довольно просто и, значит, не будут испытывать неудобств с подобными (биометрическими) системами».

Кроме того, на лице каждого человека «написаны» его история, симпатии и антипатии [Тен95], а в рамках физиогномики [Вел98, Гон97] по лицу человека можно определить, что это за человек, каково его предназначение, какими он «наделен» болезнями, в каком эмоциональном состоянии находится в настоящий момент и каковы его чувства и намерения по отношению к окружающим его людям.

Как первое, так и второе определяют практический интерес для компьютерного (автоматического) распознавания лиц вообще и решению задач идентификации и аутентификации человека по его лицу в частности. Именно эти проблемы и задачи решаются соответствующими системами в рамках биометрики и на основе биометрических характеристик лица.

**Информационные знаки лица человека.** При обычном анализе лица человека обращают на себя внимание следующие информационные знаки [Вел98]:

- форма лица (например, круглая, квадратная, треугольная...);
- соотношение частей лица между собой (лоб, средняя и нижняя части лица);
- форма лба, скуловой области и подбородка;
- форма и размер ушной раковины, способ ее прикрепления, а также форма частей ушной раковины (мочки, козелок, противокозелок);
- симметрия лица (симметрия или явная асимметрия);
- форма, величина и местоположение глаз, рта, носа;
- линии локализации морщин и т. д.

В методах автоматического распознавания лиц по видеопортретам используют эти же информационные знаки в различных их сочетаниях, что в большинстве случаев определяется тем, какой именно портрет распознается — ориентированный в фас, в профиль или одновременно два портрета: в фас и профиль. Если же говорить о коммерческих системах, то они ориентированы на распознавание лиц, представленных в фас (с возможным поворотом и с наклоном головы не более  $\pm 15^\circ$ ), в связи с чем практически не используют информационные знаки ушных раковин и профиль лица. Хотя некоторая информация о профиле все же может быть получена и из изображения лица строго в фас.

При этом два источника информационных знаков лица — глаза и ушные раковины — представляют собой самостоятельные источники биометрических характеристик человека и определяют отдельные направления в биометрике.



### 2.3.1. ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ ЛИЦ

Результат распознавания лиц автоматической системой зависит от качества исходных данных (например, фото- и видеопортретов); выбора исходных информационных знаков лица и/или способа выделения признаков из них; метода, используемого для распознавания, и критериев распознавания.

Если говорить о методах, то в технической литературе, посвященной проблеме автоматического распознавания лиц, наиболее часто ссылаются на основополагающие публикации [Kir90, Tur91, Bru93, Bel97], представляющие следующие методы:

- корреляционный метод (метод согласованной фильтрации);
- метод, основанный на преобразовании Карунена—Лоэва и понятии «Eigenface» (собственных лиц);
- метод, основанный на линейном дискриминантном анализе и понятии «Fisherface», названном так в честь Роберта Фишера, использовавшего дискриминантный анализ в распознавании образов.

Относительно простым среди перечисленных является корреляционный метод. Обычно он реализуется в пространстве исходных признаков (исходных изображений) и аналогично тому, как реализуется согласованная фильтрация. Важен сам факт существования корреляции между распознаваемым образом лица и эталоном, хранящимся в базе данных. Если условия получения новых образов соответствуют условиям получения эталона (освещение, пункт наблюдения лица, определяющий его наклон и поворот, масштаб, фон и т. д.), то корреляция между ними близка к единице, а уровень распознавания в целом доходит до 96 % [Bru93]. Однако, если фон, масштаб, повороты и наклоны изменяются, то обычная линейная корреляция не устанавливает подобия двух образов и становится практически бесполезной.

Развитием корреляционного метода является переход от исходного пространства признаков (исходного изображения) к пространству инвариантов Фурье—Меллина (элиминирующих масштаб и заменяющих поворот на циклический сдвиг), что позволяет на основе той же самой согласованной фильтрации достичь высоких оценок корреляции между двумя образами [Che94, Ben98]. Проблемой по-прежнему здесь остается высокая размерность пространства признаков, соответствующая размерам исходных изображений и, кроме того, преобразование Фурье—Меллина существенно усложняет весь вычислительный процесс согласованной фильтрации.

Метод, использующий преобразование Карунена—Лоэва, позволяет осуществить значительную редукцию исходного пространства признаков, оставляя для дальнейшего использования только те из них, которые имеют принципиальное значение для конкретного класса образов. При этом непосредственное влияние условий получения образов на его представление в редуцированном



пространстве признаков становится не так заметно, а сравнение образов в этом пространстве значительно упрощается. Уровень правильного распознавания здесь стабильно достигает 80 % даже при заметных изменениях условий получения образов лиц по сравнению с условиями получения эталона.

Метод, основанный на линейном дискриминантном анализе, так же как и предыдущий, позволяет реализовать редукцию исходного пространства признаков, при этом он существенно улучшает кластеризацию образов в редуцированном пространстве признаков. Последнее обеспечивает еще более высокий уровень распознавания, достигающий до 99 % даже при изменениях условий получения образов.

На сегодняшний день основное развитие при распознавании лиц получили методы, ориентированные на следующее:

а) репрезентативный характер исходных данных — использование для обучения системы распознавания того диапазона изменений образов лиц (освещение, ориентация, масштаб, фон и т. д.), в рамках которого система должна распознавать все другие образы;

б) уменьшение размерности исходных данных — применение редукции исходного пространства признаков с представлением исходных образов в редуцированном пространстве признаков;

в) реализацию процесса распознавания в редуцированном пространстве признаков;

г) использование процедур сопоставления распознаваемого образа с эталоном с оценкой меры подобия между ними (на основе какой-либо метрики), используемой в качестве критерия распознавания.

В связи с изложенным выше в настоящем разделе будет дан обзор систем распознавания, использующих базы, которые призваны обеспечить репрезентативность исходных данных, и систем, в которых редукция признаков основана на методах анализа главных компонент (англ. PCA), преобразовании Карунена—Лоэва (ПКЛ, англ. KLT) и линейном дискриминантном анализе (англ. LDA\*). Обоснованием для их детального представления может служить тот факт, что большинство реализованных в последнее время на практике систем распознавания людей по их лицам используют именно эти подходы и методы [Gou96, Swe96, Hon98, Lam98, Mul98, Tsa98, Swe99, Liu00, Гла97, Гла97, Сам98 и др.]. Наконец, о важности перечисленных выше методов в решении проблемы «Face Recognition» и в приложении к коммерческим биометрическим системам свидетельствует также обзор, представленный в специальном выпуске журнала «Computer» [Pen00, Отс00].

Очевидно и то, что практическое применение этих, относительно непростых (с вычислительной точки зрения), методов стало

---

\* PCA — Principal component analysis, LDA — Linear discriminant analysis, KLT — Karhunen—Loeve Transformation, ПКЛ — преобразование Карунена—Лоэва, ЛДА — линейный дискриминантный анализ.



возможным благодаря мощному развитию вычислительных средств, в том числе ресурсов персональных компьютеров, используемых в биометрических системах.

### 2.3.2. БАЗА ДАННЫХ ИЗОБРАЖЕНИЙ ЛИЦ

Если говорить о качестве исходной базы данных, то здесь подразумевается:

- ее репрезентативность;

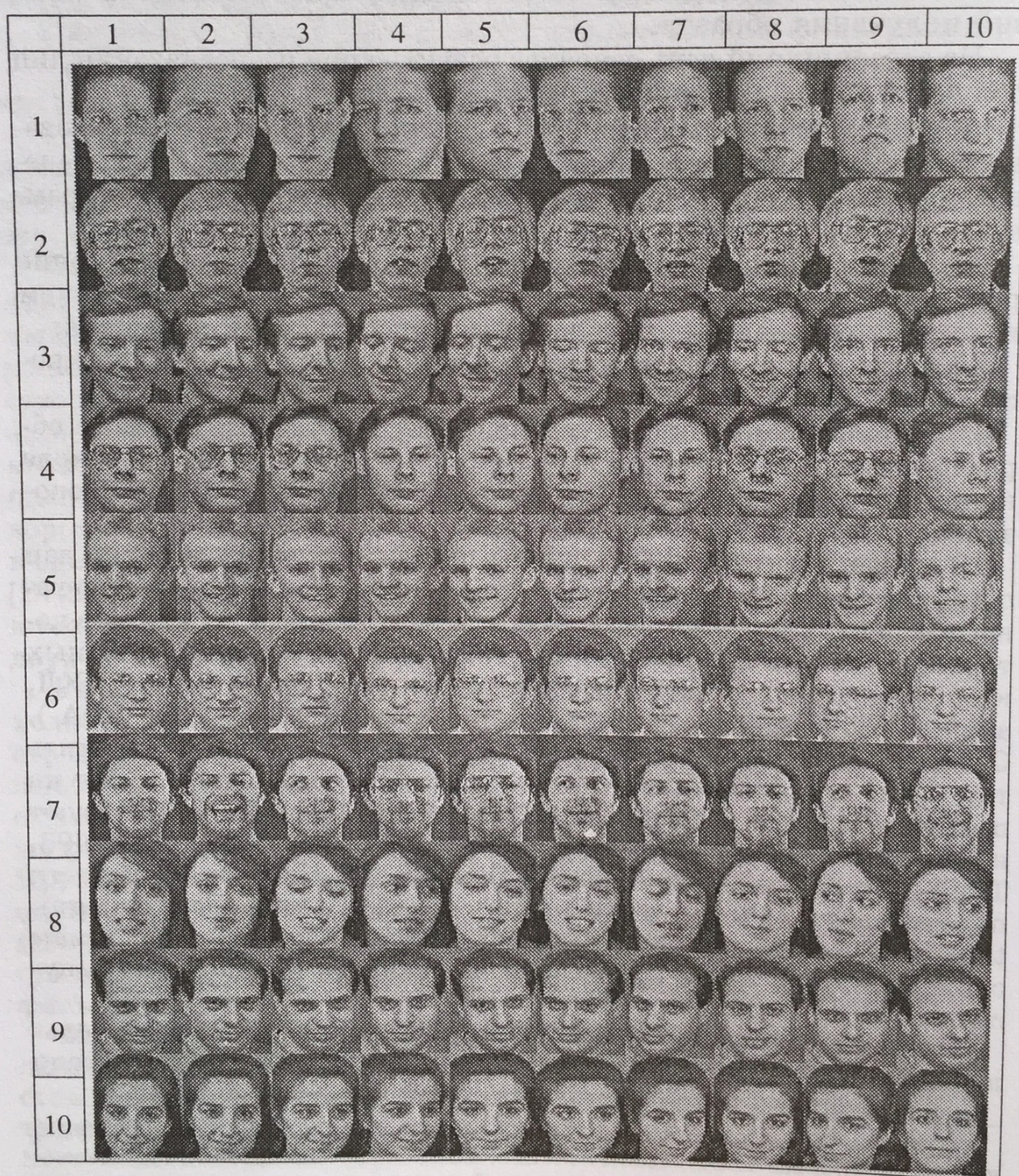


Рис. 2.3.1. Первые 10 классов образов базы данных ORL Database of Faces



- факт и/или способ структурирования данных;
- качество самого образа, в том числе:
  - размер каждого образа, выраженный, например, в числе пикселей;
  - контраст и прорисовка деталей лица;
  - фон, на котором находится основная часть лица;
  - отсутствие мешающих деталей на области лица и т. д.

В свою очередь, репрезентативность базы данных (лиц) и способ ее структурирования, выраженные, например, в числе классов и числе образов в каждом классе, определяются необходимостью наличия в базе образов лиц с различным поворотом головы — вверх/вниз, влево/вправо и их комбинаций, присутствием или отсутствием дополнительных предметов (очки, серьги...), а также различным выражением на них. Важно при этом еще и соблюдать более или менее одинаковые условия получения образов лиц (освещенность, размер самого лица на фоне всего образа, незначительные возрастные изменения лиц, представленных в каждом классе).

Довольно часто для тестовой проверки качества систем распознавания используется база данных ORL Database of Faces [Orl\*\*] как отвечающая вышеперечисленным требованиям и наиболее доступная широкому классу исследователей проблемы «Face Recognition». Первые 10 классов образов этой базы данных представлены на рис. 2.3.1. Целиком эта база данных представлена в Приложении.

ORL Database of Faces содержит 400 образов по 10 в каждом классе. Каждый образ представлен изображением размера  $112 \times 92$  пикселя в 256 уровнях градации шкалы GRAY и всегда на темном фоне. Фотографии, из которых получены образы лиц, выполнены в разное время (некоторые с разницей более двух лет). Репрезентативность данных обеспечивается некоторыми изменениями масштаба лица, угла наблюдения в момент фотографирования (что определяет изменение наклона и поворота головы) и условий освещения.

Получение более полных баз данных — процесс очень трудоемкий и длительный, хотя существуют базы данных с числом образов, превосходящим 1000. Примером здесь может служить база FERET [Phi9, Pen00], используемая для тестовой проверки коммерческих систем распознавания лиц.

### 2.3.3. ВВЕДЕНИЕ В МЕТОДЫ PCA, ПКЛ И LDA

Суть анализа главных компонент и связанного с ним ПКЛ видна из рис. 2.3.2.

Цель анализа — выявить основные изменчивости в изображениях лиц (обучающем наборе исходной базы данных) и описать эти изменчивости при помощи нескольких собственных векторов (это реализуется в блоках 3–5 на рис. 2.3.2).



Основные изменчивости представлены матрицей ковариации, порядок которой (пусть он будет равен, например,  $N$ ) соответствует размерности вектора исходных признаков, описывающих каждый образ. Собственные числа матрицы ковариации являются спектром этой матрицы в ортогональном базисе собственных векторов и диагонализуют ее. При этом основная изменчивость матрицы ковариации (а также ее основная энергия в соответствии с равенством Парсевала) описывается всего  $p$  наибольшими собственными числами, при этом  $p \leq N$ .

В дальнейших преобразованиях участвуют только эти наибольшие собственные числа и соответствующие им собственные векторы.

В свою очередь, ПКЛ, реализуемое в блоке 6, — ортогональное преобразование в базисе, сформированном из собственных векто-

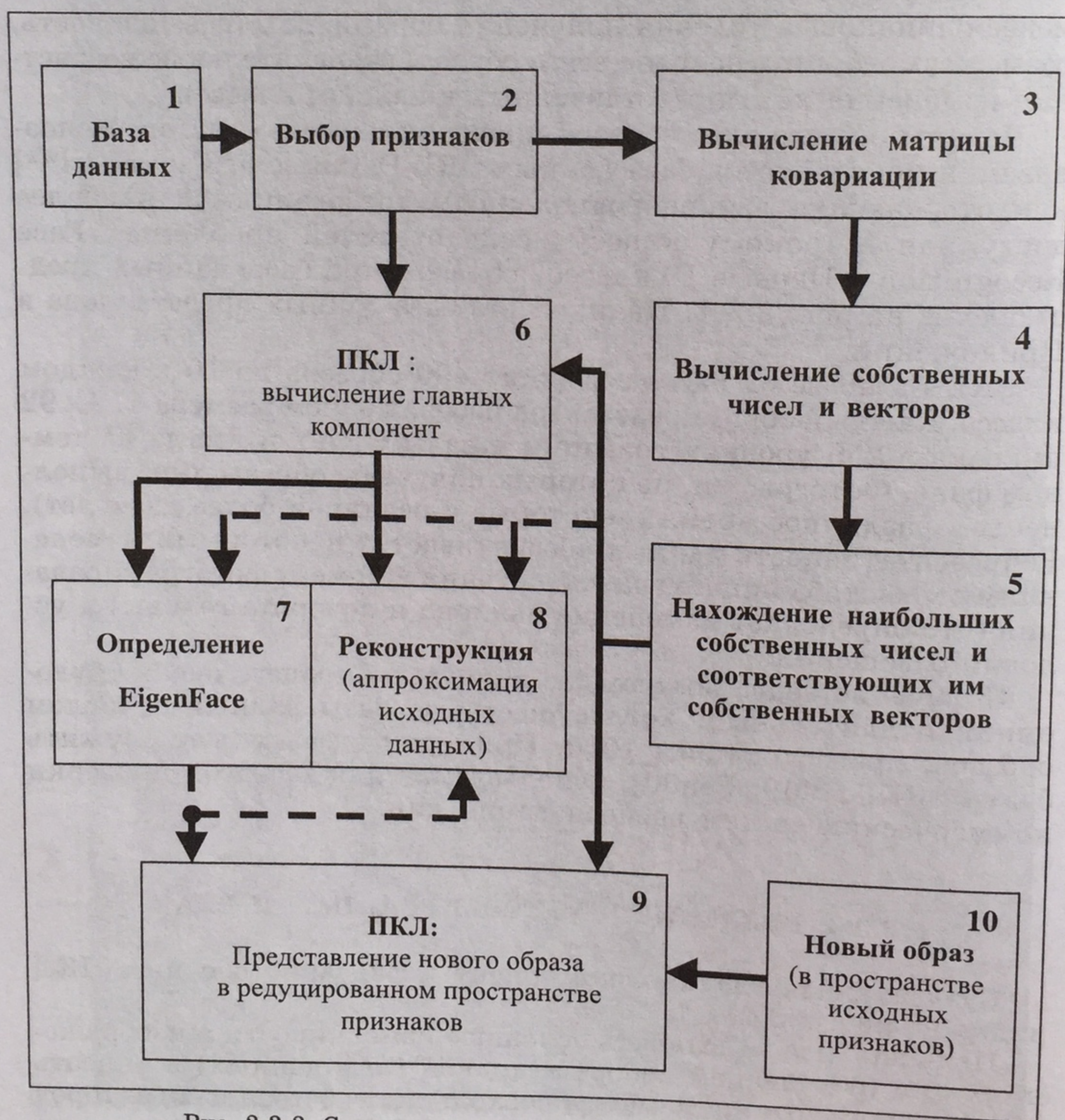


Рис. 2.3.2. Структура вычислений в методе PCA и ПКЛ



ров. Входными данными здесь являются образы, представленные в исходном признаковом пространстве размерности  $N$ , а выходными данными — главные компоненты: спектр, определяемый только  $p$  собственными векторами, выбранными выше. Таким образом, ПКЛ позволяет трансформировать  $N$ -мерное пространство исходных признаков в  $p$ -мерное пространство признаков. Этот процесс называется *редукцией пространства признаков*, и именно здесь в результате ПКЛ определяются главные компоненты.

Обратное ПКЛ для полученных значений главных компонент приводит к их представлению в форме «собственных лиц» (или, например, «эластичных моделей лиц» [Lan97]), что реализуется блоком 7. Полная реконструкция или аппроксимация образов осуществляется аналогично (блок 8).

Каждый новый образ, предназначенный для распознавания, также подвергается ПКЛ в том же базисе собственных векторов (блок 9), в результате чего получается его представление в редуцированном пространстве признаков.

Каждый образ (из базы или новый) может быть также представлен как линейная комбинация изображений «собственных лиц» или аппроксимирован через параметры соответствующих эластичных моделей лиц.

В редуцированном пространстве признаков значительно проще реализуется как решение задачи кластеризации образов, так и сравнение образов между собой, поскольку каждый из них представляется существенно меньшим числом признаков и только теми, которые описывают их главные особенности.

При использовании метода РСА сокращается пространство признаков (в нашем случае в  $N/p$  раз), однако он не всегда позволяет удовлетворительно решать задачи кластеризации.

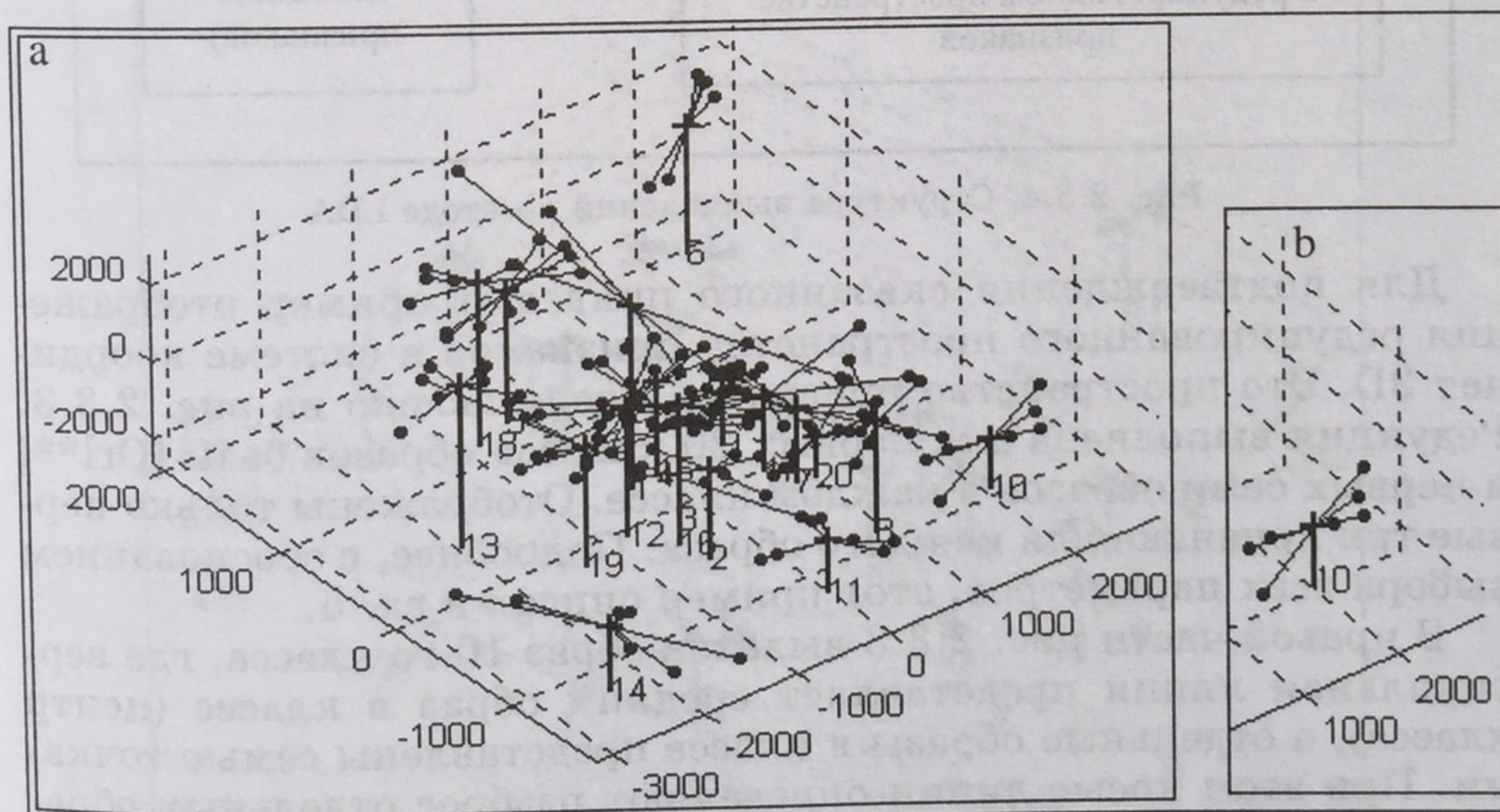


Рис. 2.3.3. Отображение редуцированного пространства признаков в форме 3D: а — редукция для первых 20 классов образов базы; б — образ 10-го класса



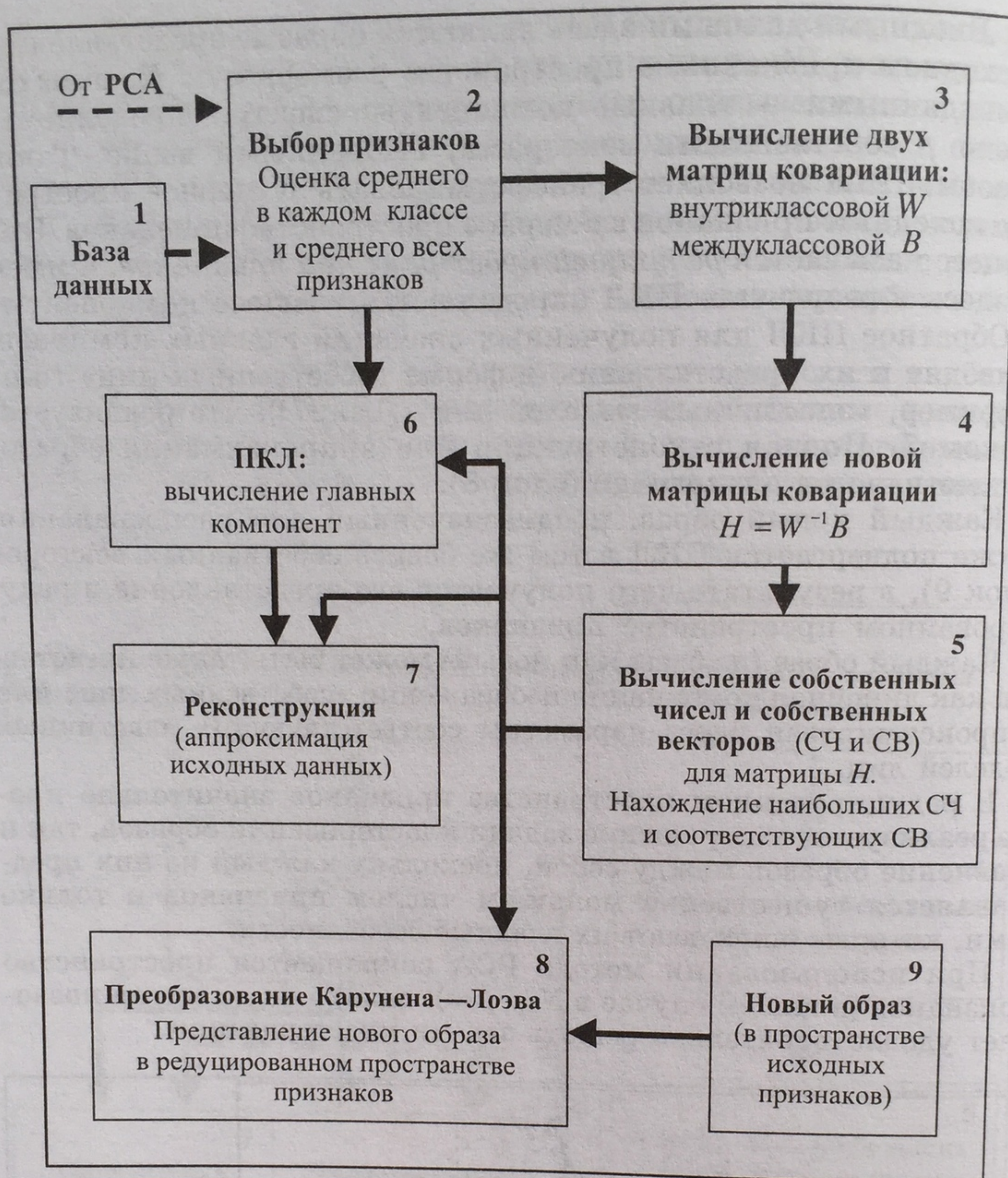


Рис. 2.3.4. Структура вычислений в методе LDA

Для подтверждения сказанного приведем пример отображения редуцированного пространства признаков в системе координат 3D. Это пространство признаков представлено на рис. 2.3.3. Редукция выполнена для первых 20 классов образов базы [Orl\*\*] и первых семи образов в каждом классе. Отображены только первые три признака для каждого образа. Подробнее, с обоснованием выбора всех параметров, этот пример описан в гл. 5.

В правой части рис. 2.3.3 выделен образ 10-го класса, где вертикальная линия представляет средний образ в классе (центр класса), а отдельные образы в классе представлены семью точками. При этом косые линии определяют разброс отдельных образов относительно среднего образа в классе (относительно центра класса).



Оценивая результат представленной кластеризации, отметим, что здесь не *максимизированы* расстояния между классами и не *минимизированы* внутриклассовые расстояния между образами в классе. Иными словами, из представленного рисунка видно, что некоторые образы значительно отходят от центров своих классов и так далеко, что попадают «на территорию» соседних классов (перекрывают и свои и чужие классы), а сами центры классов имеют не совсем равномерное распределение в пространстве признаков.

Если теперь редуцированное методом ПКЛ пространство признаков рассматривать как исходные данные, то, используя в приложении к ним метод LDA, можно создать дополнительные возможности для редукции пространства признаков и при этом существенно улучшить их кластеризацию.

Структура вычислений, реализуемая в методе LDA, показана на рис. 2.3.4, и, поскольку он имеет много общего с методом PCA, пояснений не требуется.

Результат кластеризации пространства признаков (метод PCA + LDA, реализованные в [Кис98, Кис99 и Кис00]) для базы данных, содержащей те же 20 классов образов, представлен на рис. 2.3.5.

Сравнивая результаты кластеризации на рис. 2.3.3 и 2.3.5, можно отметить, что ожидаемый эффект достигнут: отдельные образы в классах еще теснее концентрируются около своих центров, практически не «залезая на чужие территории», а сами центры отдалены друг от друга еще больше.

Представленные подходы используются сегодня при построении систем распознавания лиц [Swe96, Swe99], что обеспечивает высокие показатели распознавания как с очень большими базами

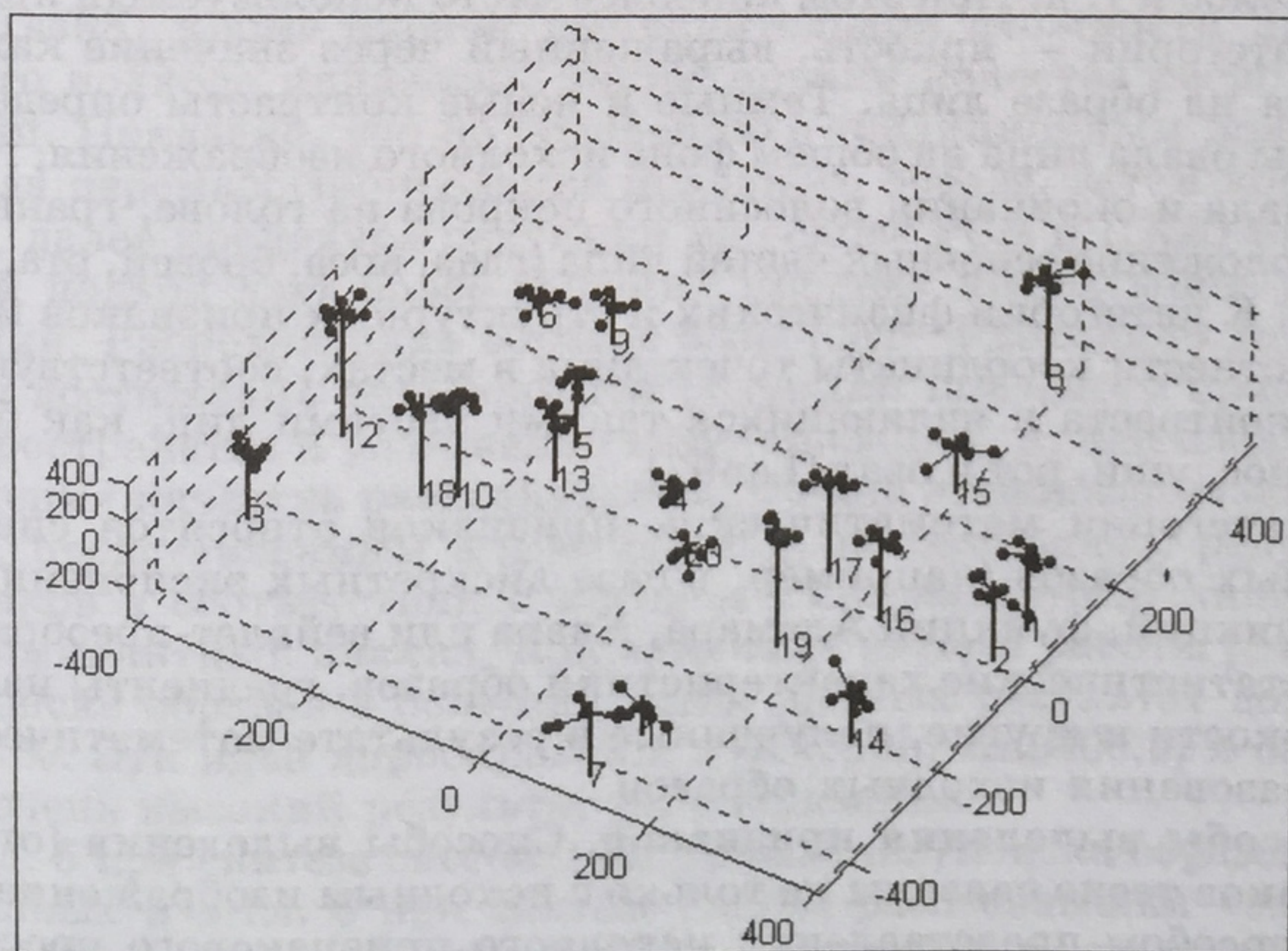


Рис. 2.3.5. Кластеризация признаков по методу PCA + LDA



данных (распознавание, поиск близких образов в базах), так и построения систем распознавания лиц для ограниченного числа классов и очень малого числа образов в классах.

Максимальное число признаков после двухэтаповой процедуры редукции не может быть больше  $(K - 1)$ , где  $K$  — число классов. Такое существенное сокращение размерности пространства признаков при сохранении главных характеристик в описании каждого отдельного образа (базы данных или нового) создает необходимые и достаточные условия для использования простых механизмов сопоставления образов в процессе их распознавания.

#### 2.3.4. ОТБОР ПРИЗНАКОВ ПРИ РАСПОЗНАВАНИИ ЛИЦ

В настоящем разделе будут рассмотрены категории признаков и способы их выделения в задачах распознавания лиц, построение исходного признакового пространства и практические способы его редукции, описанные в технической литературе по проблеме «Face Recognition».

**Категории признаков.** Как отмечено в [Tou74], выбор и выделение признаков играют в распознавании образов центральную роль. В [Tou74] определены три категории признаков: физические, структурные и математические.

В приложении к задачам распознавания людей по лицам, физические и структурные признаки образов лиц выражаются через формы лица (овал лица, геометрия его основных частей) и цвет лица, волос и т. д. При этом наиболее часто используемый признак этих категорий — яркость, выраженный через значение каждого пикселя на образе лица. Темные и ясные контрасты определяют границы овала лица на общем фоне исходного изображения, границы начала и окончания волосяного покрова на голове, границы и местоположение основных частей лица (глаз, носа, бровей, рта, ушей и т. д.). К категории физических и структурных признаков можно также отнести координаты точек лица в местах, соответствующих смене контраста и являющихся такими частями лиц, как брови, глаза, нос, уши, рот и овал [Lan97].

К категории математических признаков относятся спектры исходных образов (например, в базе дискретных экспоненциальных функций, функций Адамара, Хаара или вейвлет-преобразования), статистические характеристики образов, градиенты изменения яркости и другие, полученные в результате математического преобразования исходных образов.

**Способы выделения признаков.** Способы выделения (отбора) признаков тесно связаны не только с исходным изображением, но и со способом представления исходного признакового пространства и способом его редукции.



Яркие признаки «собирают» в вектор признаков путем конкатенации столбцов (или строк) исходного изображения лица, представляемого матрицей данных. Однако при размерах исходного изображения  $M \times N$  пикселей, вектор признаков содержит  $MN$  компонент, а матрица ковариации имеет порядок  $MN$ . Например, при значениях  $M = 112$  и  $N = 92$  этот порядок составит величину 10304. Гарантированное нахождение собственных чисел и соответствующих им собственных векторов становится в таком случае практически невозможным. Поэтому в практике распознавания лиц с использованием ярких признаков перед выполнением конкатенации исходный образ лица перемасштабируют к существенно меньшему размеру.

Оригинальный метод представления исходного пространства признаков изложен в [Tsa98]. Здесь исходный образ лица рассматривается как набор столбцов и строк, являющихся самостоятельными векторами признаков. Далее процедура PCA применяется дважды: отдельно для столбцов и отдельно для строк каждого образа. В результате редукции исходного пространства признаков получается матрица, имеющая существенно меньший порядок, чем порядок матрицы исходного образа.

Как отмечено в [Tsa98], такой подход упрощает использование процедур PCA и повышает точность вычисления собственных чисел и векторов, поскольку порядки матриц ковариаций равны размерам изображения лица. И, кроме того, предложенный способ отбора и редукции признаков делает систему распознавания более устойчивой к изменениям яркости исходных образов, а также к циклическому сдвигу собственно лица в поле исходного изображения, и к шумам, наложенным на исходный образ.

В [Tsa98] кроме оценок вычислительной сложности предложенного подхода приведены также оценки качества распознавания лиц. Показано, что по сравнению с традиционным подходом (сначала перемасштабирование исходного образа, затем конкатенация, далее процедура редукции на основе PCA и ПКЛ) предлагаемый подход дает более высокий процент правильного распознавания. Распознавание образов базы [Orl\*\*] доходило до 85 %.

Для достижения еще большей редукции исходного признакового пространства и улучшения кластеризации, и, как следствие, улучшения качества распознавания, можно воспользоваться подходом, представленным в [Swe99], т. е. по окончании редукции признаков в соответствии с методом PCA выполнить линейный дискриминантный анализ. Как отмечают авторы работы [Swe99], при поиске образов в больших базах данных результат доходил до 100 %. Эти идеи апробированы в [Kuc99b, Kuc00c,e] и обеспечили очень высокий результат распознавания.

В гл. 5 при синтезе системы распознавания/поиска образов в базах данных и в гл. 6 при синтезе систем распознавания «on-line» мы подробно рассмотрим способ отбора и редукции признаков [Tsa98], соединив его с линейным дискриминантным анализом по [Swe99].



В категории математических признаков интерес представляют спектральные, например, признаки, полученные как результат преобразования Фурье исходного образа. Процесс их выделения состоит в следующем: выполняется двумерное дискретное преобразование Фурье исходного изображения лица; вычисляются абсолютные значения полученного спектра; далее выполняется циклический сдвиг (на половину размера исходного изображения) спектра по двум осям. Во-первых, такой спектр становится инвариантом циклического сдвига образа (циклического сдвига лица на фоне его окружения), а во-вторых, симметричным (поскольку исходные данные принадлежат к данным действительного типа) и, следовательно, размер признакового пространства может быть сокращен вдвое. Кроме того, как показывают исследования, полная информация об образе лица содержится всего в 20 пространственных частотах. Таким образом, размер исходного пространства признаков для каждого образа может составить максимум  $(20 \cdot 20)/2 = 200$  элементов. Однако наш собственный опыт использования этих признаков показывает [Кис98, Кис99], что достижение высокого качества распознавания лиц возможно лишь при реализации в системе распознавания подходов, связанных с так называемыми локальными РСА. В отличие от «глобального РСА», в методе локального РСА анализ главных компонент и ПКЛ выполняется не для всей совокупности исходных данных как единого целого, а для каждого  $k$ -го класса исходных данных отдельно, где  $k = 1, 2, \dots, K$ . Иными словами, в этом методе выполняется  $K$  процедур РСА, где  $K$  — число классов образов в исходных данных. Распознавание каждого нового образа выполняется также  $K$  раз и заканчивается поиском, например, минимума расстояния из  $K$  вычисленных расстояний. Использование метода локальных РСА подробно описано, например, в [Mul98].

Рассмотрим еще один пример использования спектральных признаков. Как упоминалось выше, применяется также переход от пространства исходных спектральных признаков к пространству инвариантов Фурье—Меллина, элиминирующих масштаб и заменяющих поворот в плоскости  $XU$  на циклический сдвиг. Переход реализуется преобразованием спектра мощности (исходного изображения) из декартовой системы координат в полярно-логарифмическую [Che94, Ben98] или полярно-параметрическую систему координат [Кис97а, Кис98а,с]. Проблемой по-прежнему здесь остается высокая размерность исходного пространства признаков, соответствующая размерам исходных изображений, и, кроме того, преобразование Фурье—Меллина несколько усложняет весь вычислительный процесс в задачах распознавания [Кис98b, Кух99b, Кух99с]. Для сокращения пространства признаков рассматривается только половина исходного спектра, из которого используется ограниченное число пространственных частот.

Дру  
веден в  
пиксел  
ный ве  
25 мас  
знаки  
дой ма  
жений  
ве мето  
Резу  
соком  
для это  
познава  
и учет  
следних  
ственно  
дится п  
наковун

Оста  
ющемся  
В да  
области



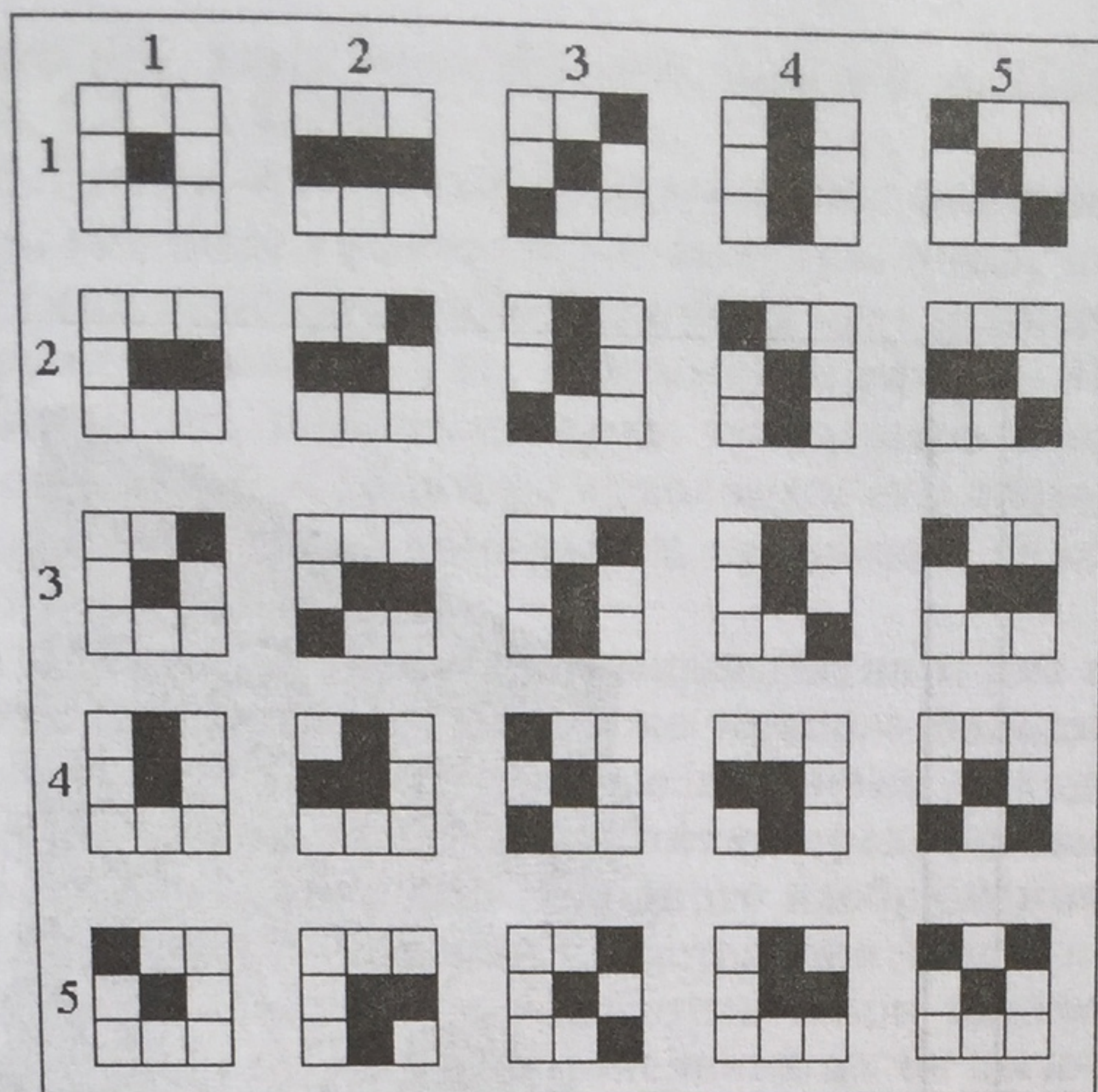


Рис. 2.3.6. Маски для отбора признаков из образов лиц по методу [Gou96]

Другой пример из категории математических признаков приведен в [Gou96]. Здесь предлагается 25 масок (размеров  $3 \times 3$ ,  $5 \times 5$  пикселей и т. д.), с помощью которых вычисляется 25-компонентный вектор признаков для каждого исходного изображения. Вид 25 масок размером  $3 \times 3$  пикселя представлен на рис. 2.3.6. Признаки вычисляются путем сканирования образа поочередно каждой маской (аналогично методам локальной обработки изображений [Прэ82]), а кластеризация признаков реализована на основе метода LDA.

Результаты, представленные в [Gou96], свидетельствуют о высоком качестве распознавания лиц в системе «on-line», однако для этого требуется большая база лиц для синтеза системы распознавания (от 50 до 100 образов на класс и не менее 100 классов) и учет специфических требований к самим образам. В числе последних — определенный тип лица и определенный размер собственно лица в поле изображения, черный фон, на котором находится портрет, распознаваемый человек должен быть одет в одинаковую одежду и др.

### 2.3.5. АНАЛИЗ ЛОКАЛЬНЫХ ПРИЗНАКОВ

Остановимся еще на одном подходе — распознавании, базирующемся на анализе локальных признаков лица [Pen00].

В данном подходе на исходном изображении локализуются области, соответствующие, во-первых, лицу и далее на каждом



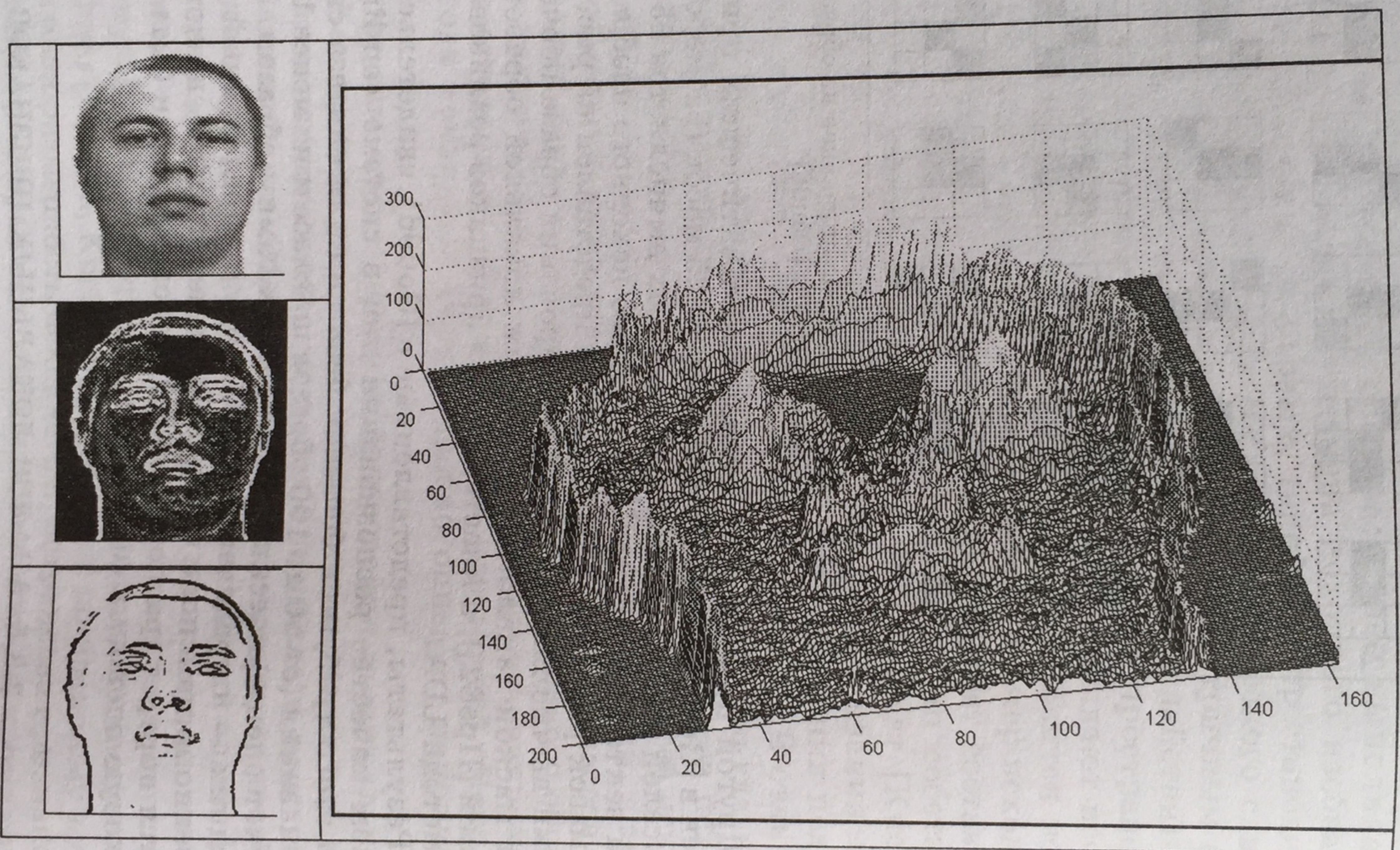


Рис. 2.3.7. Исходное изображение лица и его градиенты яркости

лице — об-  
стрируют  
На рис.  
ражение л  
фильтрац  
вой части  
ром отчет  
границы с  
Весь п  
рисунок 2  
На пол  
(точечных  
Все эти п  
ражения.  
наты  $x_1, y$   
На рис.  
показыва  
вольно хор  
нений.

Теперь  
ны рис. 2  
них показ  
поля 4. И

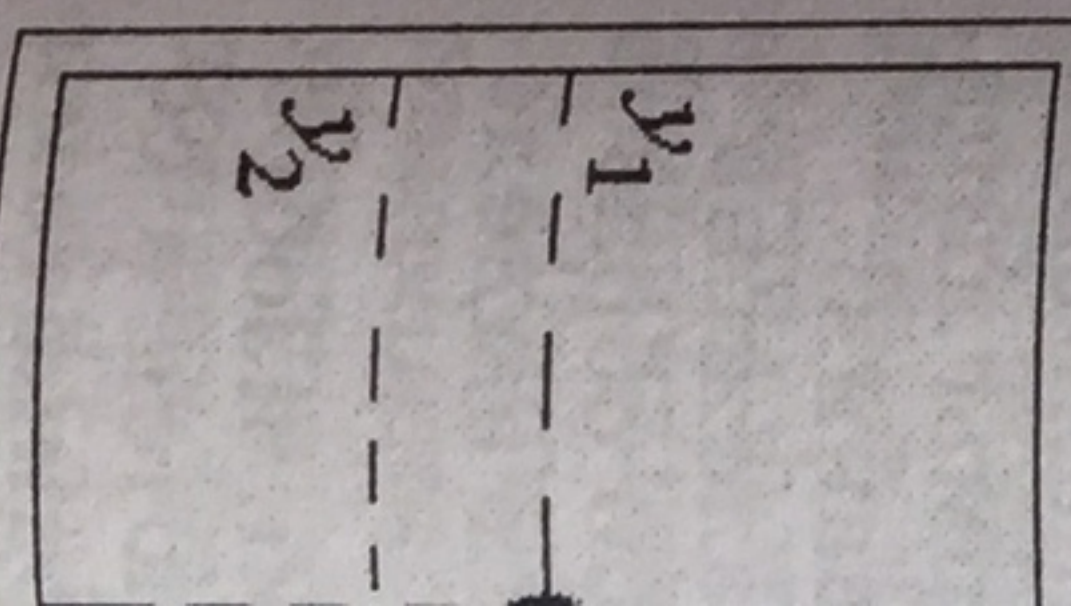


Рис.  
1 — исходное  
длящее резуль  
резул



лице — области рта, глаз, носа, бровей, щек и т. д. Подход иллюстрируют рис. 2.3.7 – 2.3.9.

На рис. 2.3.7 в левой колонке представлены: полутоновое изображение лица, его поле градиентов и контуры лица, полученные фильтрацией (или пороговой обработкой) поля градиентов. В правой части рисунка показано поле градиентов в форме 3D, на котором отчетливо видно, что максимумы градиентов очерчивают и границы области лица и границы отдельных его частей.

Весь процесс получения локальных признаков демонстрирует рисунок 2.3.8.

На поле 4 видны два больших темных пятна и два маленьких (точечных) пятна, находящихся ниже правого большого пятна. Все эти пятна определяют локальные признаки исходного изображения. В данном случае большие пятна представляют координаты  $x_1, y_1$  и  $x_2, y_2$  центров лиц исходного изображения.

На рис. 2.3.9 (с левой стороны) представлено поле 4 в форме 3D, показывающее, что максимумы, соответствующие центрам лиц, довольно хорошо локализованы и определение их не вызовет затруднений.

Теперь рассмотрим точечные пятна на поле 4. С правой стороны рис. 2.3.9 представлены еще два изображения. На одном из них показана увеличенная область правого локального признака поля 4. И на ней уже очень отчетливо видны еще два точечных

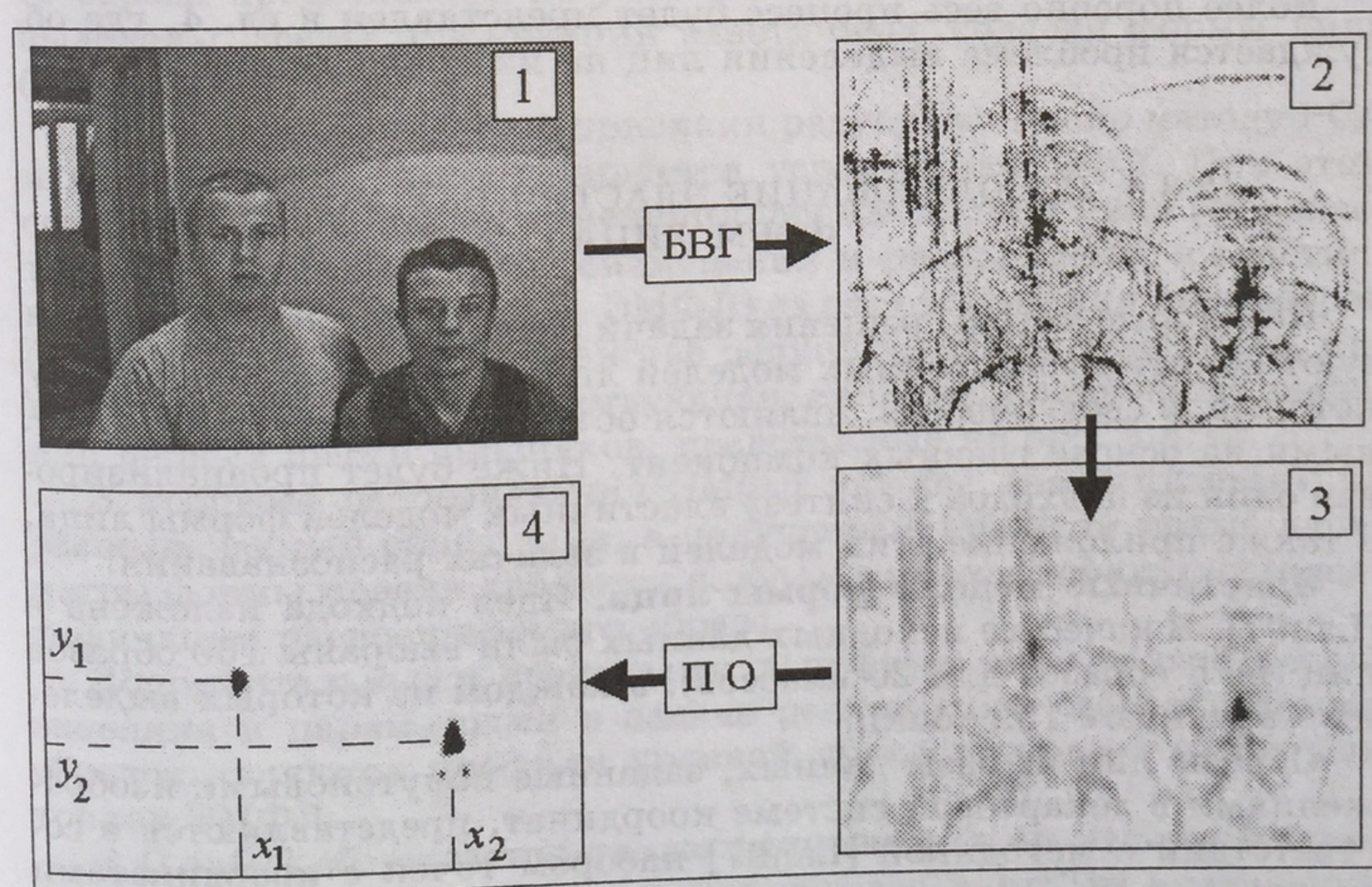


Рис. 2.3.8. Иллюстрация к методу анализа локальных признаков:  
1 — исходное изображение; БВГ — блок вычисления градиентов яркости; 2 — поле, представляющее результат анализа градиентов яркости и накопления информации о центрах лиц; 3 — результат фильтрации поля 2; 4 — результат пороговой обработки поля 3



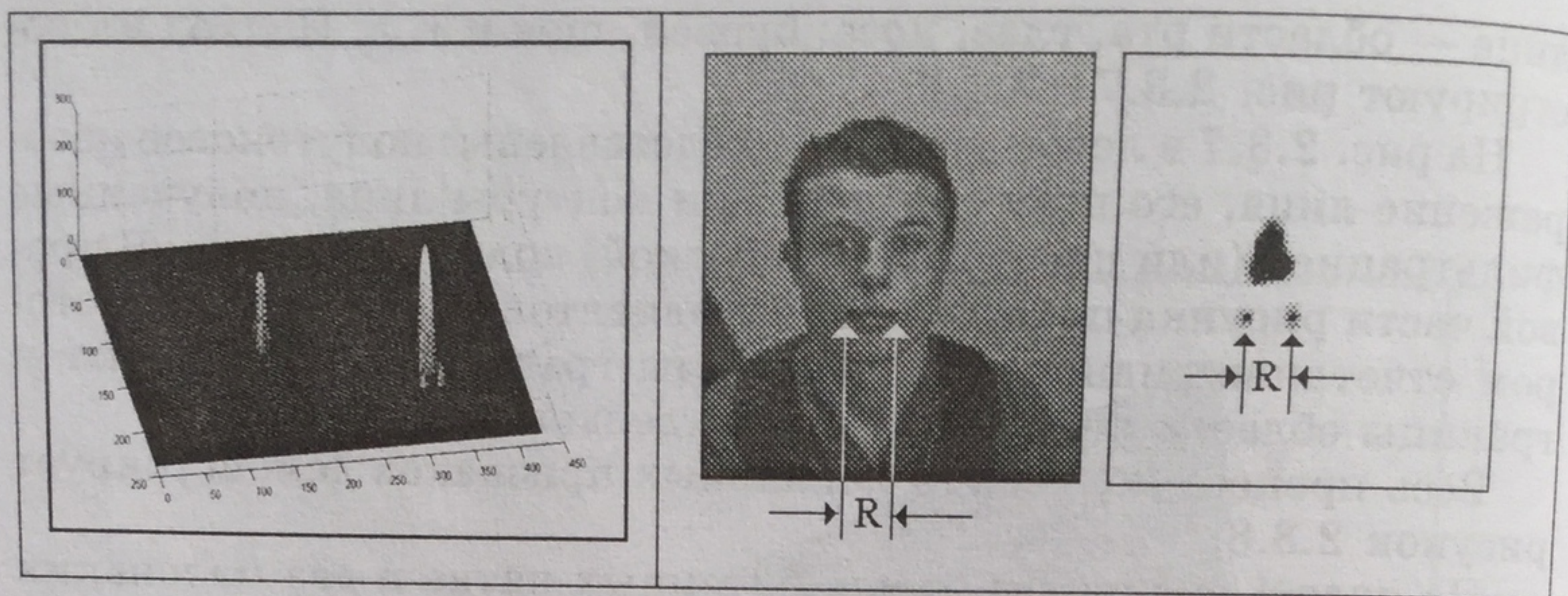


Рис. 2.3.9. Локальные признаки центров лиц в форме 3D (слева) и локальные признаки координат уголков рта

локальных признака, соответствующие, скорее всего, координатам уголков рта на правом изображении (по отношению к исходному) лица. Другие локальные признаки (например, от глаз и бровей) также просматриваются на поле 2. В целом же и все остальные локальные признаки легко обнаружить, применяя описанные подходы непосредственно к изображению самого лица.

Координаты локальных признаков и вычисленные расстояния между ними позволяют описать лицо с помощью его антропометрических точек и параметров, которые будут использованы далее при распознавании.

Более подробно весь процесс будет представлен в гл. 4, где обсуждается проблема выделения лиц на изображениях сцен.

### 2.3.6. ИСПОЛЬЗОВАНИЕ ЭЛАСТИЧНЫХ МОДЕЛЕЙ ФОРМ ЛИЦА

Некоторые методы решения задачи распознавания лиц используют понятия «эластичных моделей лица» (англ. Flexible Model), которые, в свою очередь, являются объектами, реконструированными на основе главных компонент. Ниже будет проанализирован один из подходов к синтезу эластичных моделей формы лица, а также приложение этих моделей в задачах распознавания.

**Эластичные модели формы лица.** Идея подхода изложена в [Lan97]. В качестве исходных данных были выбраны 160 образов лиц (по 8 образов для 20 классов), в каждом из которых выделены «вручную» 152 точки.

Образы лиц из базы данных, заданные полутоновыми изображениями в декартовой системе координат, представляются в соответствии с методикой [Lan97] набором точек с координатами по оси  $X$  и  $Y$ . Пример аналогичного [Lan97] выбора 152 точек приведен на рис. 2.3.10 с левой стороны.

Точки выбираются из нижней 2/3 каждого образа лица в местах, соответствующих смене контраста (яркости) и являющихся



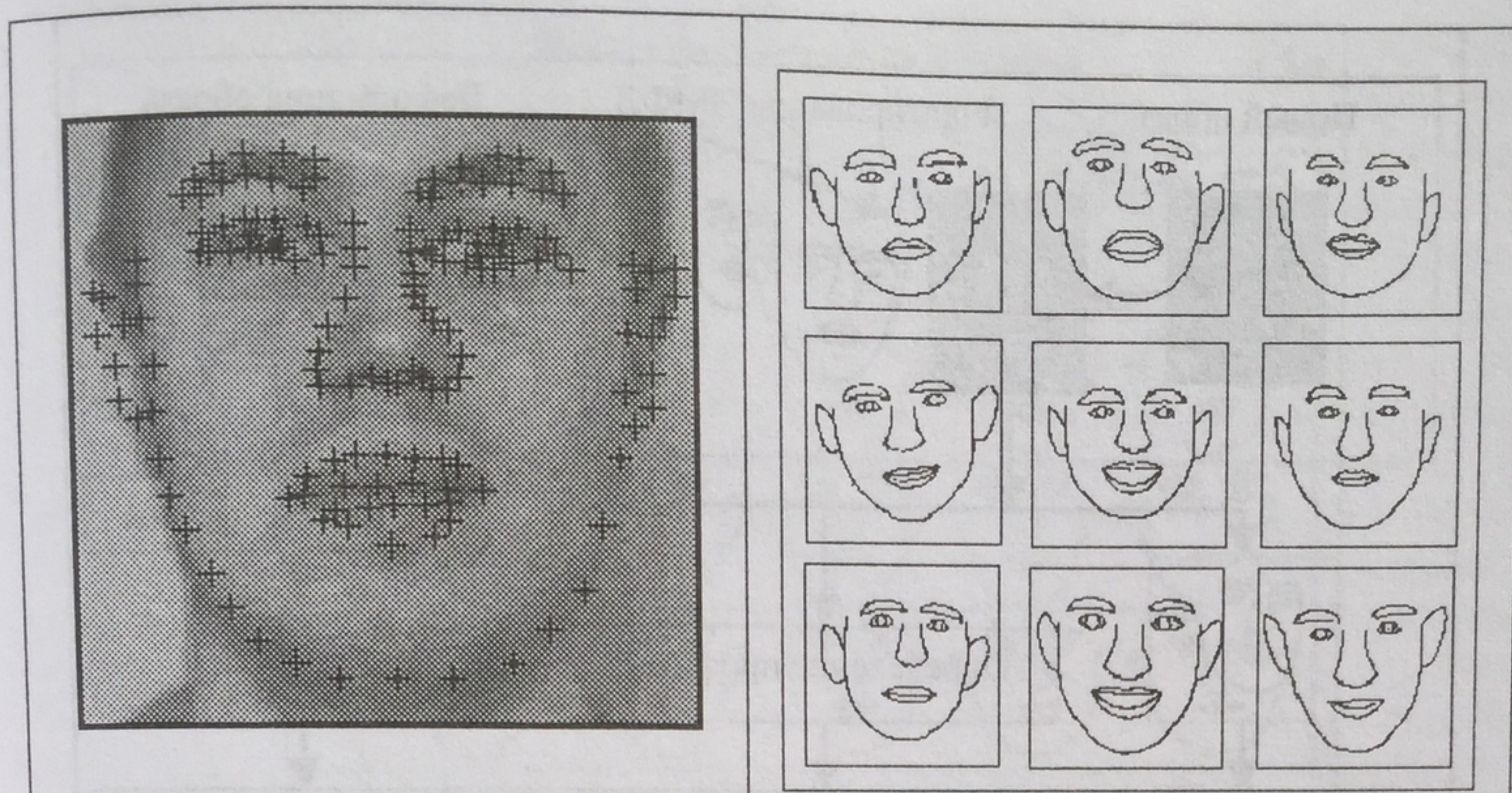


Рис. 2.3.10. Расположение точек исходной модели формы лица и формы лиц из [Lan97]

такими частями лиц, как брови, глаза, нос, уши, рот и овал лица. Эти наборы точек — первоначальные признаки — представляют собой исходные модели формы лица (МФЛ). Получаемые при этом исходные МФЛ представлены на рис. 2.3.10 с правой стороны.

Схема вычисления совокупных параметров для каждого нового образа представлена на рис. 2.3.11. В нижней части рисунка выделены сферы приложений эластичных моделей формы лица (ЭМФЛ).

Далее первоначальные признаки редуцируются по методу РСА и на их основе реконструируется усредненная МФЛ. При этом создается возможность изменения формы модели лица (ее «эластичного» растяжения/видоизменения и превращения в эластичную модель формы лица — ЭМФЛ) за счет изменения параметров (в пределах  $\pm 3\sigma$ ) в пространстве редуцированных признаков. Совокупность ЭМФЛ, соответствующая, например, изменению значений первых шести признаков, представлена на рис. 2.3.12.

В процессе распознавания ЭМФЛ как бы «натягивается» на распознаваемый образ лица, а полученные при этом новые параметры формы модели являются ключом для классификации/идентификации распознаваемого образа.

Дополнительными средствами верификации в задаче распознавания и параметрами в задаче регистрации распознаваемых образов являются профили уровней яркости образов в узловых точках ЭМФЛ.

В [Lan97] обсуждается также решение задачи реконструкции образов и построения виртуальных образов лиц. Для ее решения используется процедура деформации заданного образа (с учетом параметров, определенных по ЭМФЛ) с целью приведения его к такому положению, когда лицо повернуто строго в фас.



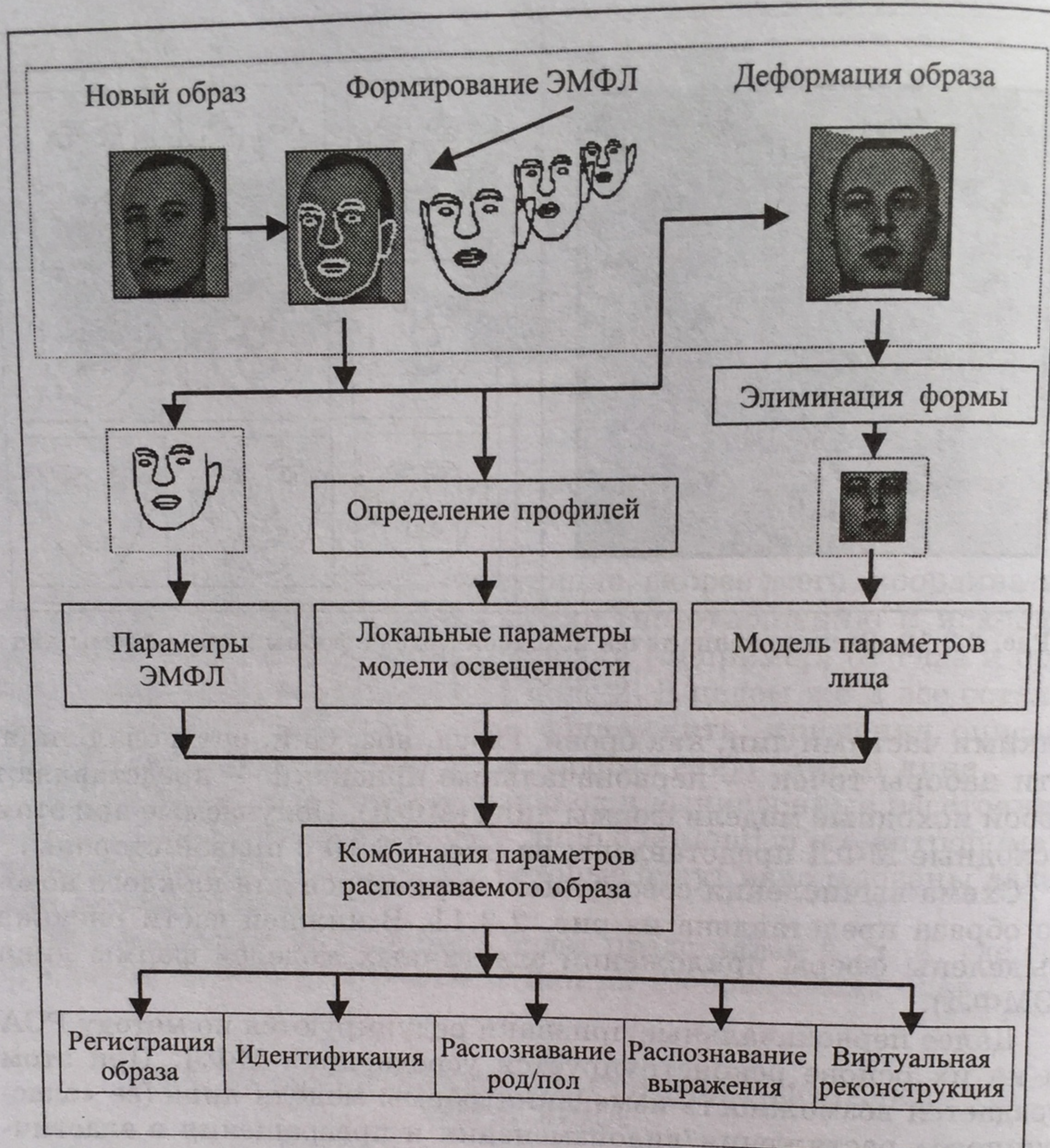


Рис. 2.3.11. Схема вычисления параметров ЭМФЛ из [Lan97]

Далее из деформированного образа выделяется нижняя часть лица, которая затем представляется моделью изменения уровней в шкале полутоновых изображений. При этом набор параметров модели описывает условия освещенности лица, дает определенную информацию о наличии бороды и усов, а также о выражении лица. Такая модель служит основой для распознавания и идентификации заданного образа в классе образов, независимых от параметров ЭМФЛ, и является основой для синтеза нового (виртуального) образа лица. При этом можно «удалить» очки и другие объекты, заслоняющие лицо, изменить выражение на лице (гримасу) и т. д.

Комбинируя все три возможности, авторы работы [Lan97] утверждают, что достигли уровня 86 %, а в некоторых случаях и 97 % классификации на тестах из 100 новых образов и 300 новых образов соответственно (по 10 образов в каждом классе).





Рис. 2.3.12. Влияние изменения параметров признаков на ЭМФЛ [Lan97]

Также в [Lan97] представлены результаты регистрации положения лица (головы в пространстве 3D) с точностью от 77 до 100 % и выражений на лице (включая агрессивность, нейтральность, радость, удивление, грусть, испуг и т. д. ) с точностью до 83 %.

### 2.3.7. КОММЕРЧЕСКИЕ СИСТЕМЫ

Среди коммерческих продуктов можно выделить системы распознавания лиц, разработанные в компаниях Visionic, Viisage, Miros, Dialog Communication Systems AG, информацию о которых можно найти на сайте [Bio\*\*] или на сайтах этих компаний: [www.faceit.com](http://www.faceit.com), [www.viisage.com](http://www.viisage.com), [www.miros.com](http://www.miros.com), [www.bioid.com](http://www.bioid.com).

Компания «Visionic» предлагает свою систему «Faceit», распознавание в которой реализовано на алгоритмах анализа локальных признаков. При этом главные из них — координаты центров глаз, относительно которых затем вычисляются и другие параметры лица, в том числе масштабный коэффициент, по которому нормируется изображение лица для приведения его к некоторой канонической форме.

В целом методы распознавания, использованные в системе «Faceit», аналогичны тем, что были представлены выше. По оценкам разработчиков, система «Faceit» может иметь широкое при-



менение, поскольку фиксирует лицо человека на значительном расстоянии от камеры и при приближении к ней, поддерживает поиск в базах данных. Известно применение системы «Faceit» для идентификации преступников по видеоданным, получаемым со 144 камер, установленных в центре Лондона. Применение системы возможно также в аэропортах и на станциях метрополитена, в школах, казино, на входе в производственные комплексы, фабрики и резиденции, а также в других местах перемещения и скопления людей.

Компания «Visage» предлагает систему «Photobook», в которой использованы методы, основанные на «собственных лицах» (Eigenface) и реконструкции образов на их основе. Иными словами, здесь реализованы алгоритмы сравнения каждого нового образа с эталоном в редуцированном пространстве признаков, а редукция выполнена с использованием PCA. Эти алгоритмы также были обсуждены нами выше. Система «Photobook» ориентирована на поиск образов лиц (близких заданному) в больших базах данных, для целей упорядочивания данных в базах и использования в системах аутентификации человека на основе образа лица с камеры и фотографии с какого-либо документа.

Компания «Miros» предлагает свою систему «TrueFace», распознавание в которой реализовано на нейронных сетях, а общая идея обработки состоит в выделении неизменяемых частей лица и их распознавании при игнорировании остальных его деталей. Демонстрационная версия системы «TrueFace» доступна через Интернет. Экспресс-результат ее тестирования представлен ниже.

Демонстрационная версия системы «TrueFace» не требует каких-либо специальных и серьезных настроек, проста в эксплуатации и показала высокое качество распознавания пользователей как в режиме аутентификации, так и идентификации. Обучение системы (учитывая то, что она построена на нейронной сети) возможно в момент регистрации пользователя и длится не более 10 секунд. Данная версия распознает более 20 пользователей.

Демонстрационная версия системы «TrueFace» имеет следующие режимы:

- настройка параметров системы;
- регистрация пользователей (до 22 человек);
- верификация пользователя системы (реализуется проверкой соответствия изображения лица, полученного с камеры, выбранному образу лица среди зарегистрированных пользователей);
- идентификация нового лица (реализуется проверкой соответствия нового изображения лица, полученного с камеры, одному из лиц базы данных, записанной в память);
- работа с базой пользователей и т. д.

Информация о пользователях (3–4 снимка на каждого с возможным изменением масштаба) хранится в базе данных в форме цветных изображений. В режиме верификации для выбора дос-



тупно только одно изображение лица конкретного пользователя. При этом выбор — это только обозначение того пользователя, с которым будет выполняться верификация. Идентификация (или верификация) осуществляется автоматически сразу же после запуска соответствующего режима и включает: захват изображения лица; слежение за его движением; печатание под распознаваемым изображением специального табло с именем распознанного пользователя, если распознавание прошло успешно.

Простые попытки обмана системы (замена «живого» лица черно-белой фотографией, например) не увенчались успехом. Ложная акцептация не выступила ни разу даже при средних значениях порогов акцептации (принятия решений). Ложный отказ наступал в режиме верификации в случаях «маскировки» лица очками, кепкой и т. д. и если в базе пользователь был зарегистрирован без аналогичных атрибутов. Результат распознавания в обоих режимах не был успешным, если существенно изменялось освещение или фон/текстура сцены по сравнению с теми условиями, которые были в момент регистрации пользователя.

Компания DCS AG представила систему «BioID» — мультимодальную биометрическую систему идентификации. Система реализует свою задачу на основе трех характеристик человека: изображения лица, голоса и движения губ [Fri00]. Именно использование последней характеристики (совместно с соответствующим динамическим эталоном) обеспечивает высокую степень защиты системы «BioID» от обмана с помощью муляжа или фотографии лица. В проспекте системы «BioID» упоминается даже о тройной защите (!?). Совместное же использование трех биометрических характеристик обеспечивает высокую точность распознавания. «BioID» рассчитана на использование в системах контроля доступа к компьютерным сетям непосредственно на рабочих местах или через терминалы удаленных пользователей. Алгоритмы, с помощью которых решается задача распознавания лиц в «BioID», основаны на моделях лиц, их антропометрических характеристиках и сравнении с эталоном из базы данных. На каждого зарегистрированного пользователя в базе данных системы можно иметь до восьми образов-эталонов. Система начинает работать с последовательностью видеоданных и переходит в активный режим с момента распознавания открытых глаз на одном из первых изображений лица. Далее происходит захват изображений лиц, их предобработка (в том числе нормализация шкалы и поворотов) и ортогонализация. После чего для каждого выделенного таким образом лица формируется вектор признаков, который сравнивается со всеми эталонами каждого класса.

А теперь обратим внимание на точность автоматического распознавания лиц.

Напомним, что изменение освещения и изменение положения головы перед камерой, смена камеры, а также существенная разница во времени при получении каждого нового изображения



значительно влияют на качество распознавания. Из [Pen00] следует, «что для фронтальных изображений лиц, сделанных в один и тот же день, приемлемая точность распознавания составляет, как правило, 95 %. Для изображений лиц, сделанных разными камерами, и при относительно разном освещении, точность распознавания снижается до 90 %. Для изображений лиц, сделанных с разницей в год, точность распознавания составляет примерно 50 % ». При этом следует отметить, что в перечисленных вариантах получения изображений лиц показатель качества распознавания FAR практически оставался неизменным (на уровне 2 %), в то время как FRR изменяется от 0,4 до 43 % [Pen00].

Из этого, в частности, следует, что, представляя качественные характеристики систем распознавания лиц, необходимо также соблюдать и те условия, при которых получены соответствующие показатели качества. И только под таким углом зрения оценивать официально публикуемые данные о точности (качестве) распознавания каждой системы.

В заключение раздела приведем два новых примера использования коммерческих систем распознавания лиц.

Для начала отметим случай практического применения систем распознавания лиц в следственной практике, представленный на сайтах [Bio\*\*].

Итак, некий шериф из Лос-Анджелеса (США) разыскивал преступника по фотороботу и для этого использовал систему распознавания лиц «FaceID» компании «ImageWare». Именно благодаря этой системе, он довольно быстро нашел фотографию человека, подозреваемого в преступлении. Информация об этом случае, представленная на сайте, сопровождалась двумя портретами — на одном из них был представлен фоторобот подозреваемого в преступлении, а на другом — фотография человека, якобы найденного в базе данных системы «FaceID».

На рис. 2.3.13 представлены эти два портрета: 1 — фоторобот, а 2 — лицо из базы данных системы «FaceID». Действительно ли эти лица принадлежат одному и тому же человеку, невозможно ни подтвердить, ни опровергнуть. Но то, что система «FaceID» выполнила свою функцию: нашла близкий фотороботу образ (и, возможно, при ограниченной базе данных преступников) — это факт!

Задолго до начала Олимпийских игр 2000 г., в аэропортах Канберры и Брисбена (Австралия) прошла испытания система паспортного контроля компании CPS. При успехе подобные системы предполагалось установить в международных аэропортах и морских портах по всей Австралии на период проведения Игр-2000. Система паспортного контроля решала две основные задачи: проводила сравнение фотографии на паспорте с изображением лица, получаемым камерой, и установление меры подобия между ними; проверку данных предъявленного паспорта по информационной сети Интерпола. Если устанавливалась высокая мера подобия лица на фотографии паспорта с лицом на изображении с камеры и по



Рис. 2.

линии Ин  
порта мог  
Решая  
дей — сра  
с изображ  
система к  
более 10 с  
между дв  
торая раз  
ственная

Кроме  
ваемого к  
ция чело  
жения, на  
развлечен

Во-пер  
числе с о  
интеллек  
ной этим  
знаков ли  
верно оце  
ных таки  
содержать  
ляющие в  
и его типо  
оценки по



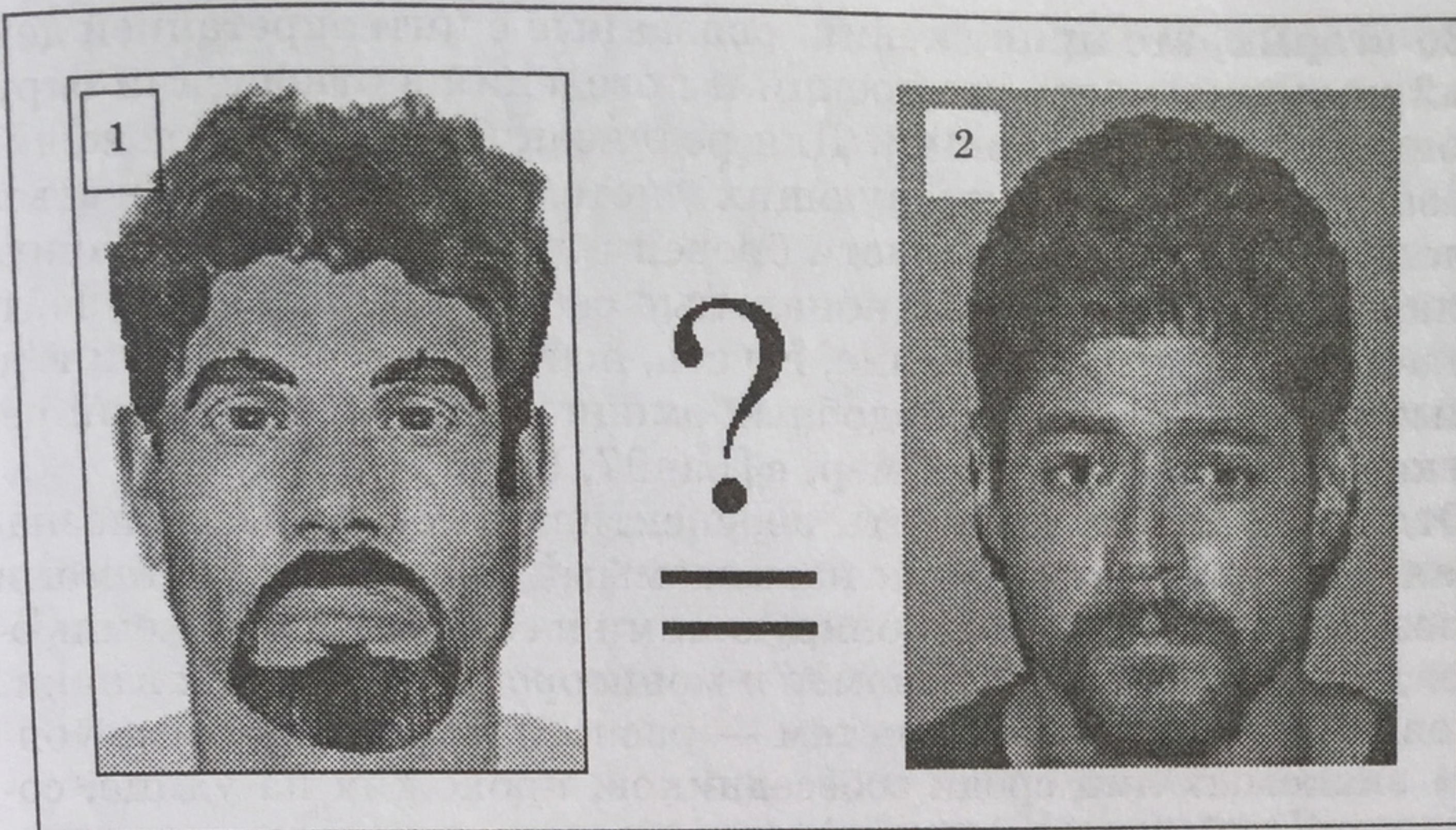


Рис. 2.3.13. Результат поиска образа, подобного заданному фотороботу

линии Интерпола не возникало никаких проблем, то владелец паспорта мог пересекать границу.

Решая одну из наиболее сложных задач распознавания людей — сравнение лица на фотографии (возможно, очень давней) с изображением лица, получаемым камерой в момент контроля, система контроля затрачивала на проверку каждого паспорта не более 10 секунд. Сложность этой задачи заключается в том, что между двумя сравниваемыми изображениями всегда была некоторая разница во времени их получения и, возможно, очень существенная разница в условиях их получения.

#### 2.3.8. НЕСКОЛЬКО СЛОВ О БУДУЩЕМ

Кроме перечисленных выше приложений систем рассматриваемого класса, главное назначение которых — это идентификация человека по его лицу, обратим внимание на два новых приложения, наиболее интересных для будущего медицины, индустрии развлечений и организации коллективных работ.

Во-первых, это приложения, связанные с физиогномикой, в том числе с оценкой физического, психического, эмоционального и интеллектуального состояния человека. В литературе, посвященной этим проблемам, описано множество общих и локальных признаков лица, а также методик его анализа, позволяющих достоверно оценить эти состояния [Теп95, Вел98, Гон97]. В базах данных таких систем (в отличие от описанных выше) должны содержаться специальные информационные комплексы, представляющие в качестве эталонов анатомические характеристики лица и его типологию, другие информационные знаки лица и методики оценки по ним состояний человека.



Во-вторых, это приложения, связанные с интерпретацией действий человека — его намерений и поведения в отношении окружающих его людей [Pen00]. Для реализации таких приложений в базах данных соответствующих систем должны содержаться эталоны лиц (особенно области бровей и глаз, области рта), определяющих различные эмоциональные состояния человека: нейтральность, радость, удивление, грусть, испуг, агрессивность и т. д. Попытки распознавания подобных эмоциональных состояний человека можно найти, например, в [Lan97, Ess97, Don99].

Отдельно можно выделить перспективные системы распознавания лиц, связанные с так называемыми носимыми системами распознавания, т. е. функционирующими на базе носимых компьютеров, миниатюрных видеокамер и мониторов, встроенных в очки. Основная функция таких систем — распознавание в режиме «on-line» знакомых лиц среди собеседников, прохожих на улице, сотрудников (в больших фирмах) или сослуживцев (например, в армии), случайно или специально попавших в поле зрения. Примеры таких систем с подробным представлением их функций уже можно найти в [Pen00].

**Идентификация человека на основе термограммы лица.** Все представленные выше системы используют биометрические характеристики лица, полученные из его двухмерного представления. Системы эти просты в эксплуатации, практически не требуют специального оборудования (использование видеокамер, как правило камер для видеоконференций, сегодня не проблема, поскольку они входят в комплект мультимедийных средств компьютера), обладают высоким быстродействием. Однако эти системы не защищены от несанкционированного доступа, поскольку могут быть «обмануты» плоскими изображениями лиц оригиналов (или фотомуляжей).

Несанкционированный доступ к системе идентификации человека по лицу становится практически невозможным, если вместо видеообраза использовать термографический образ лица. Информационными признаками такого образа являются подкожные рисунки артерий и вен, повторить которые в форме муляжа практически невозможно. Термограмма лица — уникальная характеристика человека. Для получения исходного термографического образа используются специальные камеры, улавливающие инфракрасное излучение от лица человека. Камеры могут работать в полной темноте, а информационные признаки не зависят ни от температуры самого лица, ни от выполненных пластических операций, ни от естественного старения человека. Инфракрасная камера дает приемлемый для практических приложений термографический образ лица даже в случае, когда человек находится на значительном удалении от камеры.

Следует, однако, отметить, что цена таких камер значительно выше цены (на один-два порядка) камер CCD, используемых для получения видеообразов.

О  
град  
мах  
деж  
точ  
ных  
терм  
ных  
нец  
О  
раж  
ем л  
соот  
лиц  
П  
века  
Syste  
  
Ра  
чел  
пери  
до ко  
либо  
дужн  
новы  
лочк  
при  
пред  
В бис  
ется  
некот  
Ме  
за. К  
по ра  
лича  
дужк  
В  
посре  
во вто  
ветств  
Neg00  
В с  
ставле



Очевидно, что системы идентификации, использующие термографический образ лица, могут широко использоваться в системах контроля доступа с высокой степенью ответственности и надежности. Подобные системы «обеспечивают близкую к 100 % точность распознавания независимо от использования специальных масок или даже проведения пластических операций, так как термограмма — это схема расположения внутренних кровеносных сосудов. Исследования показали, что даже однояйцевые близнецы имеют свои уникальные термограммы» [Евс97].

Однако, как и в случае использования технологии видеоизображений лиц, здесь существует проблема, связанная с положением лица перед камерой. Для получения рисунка артерий и вен, соответствующих положению лица в фас (или, например, эталону), лицо должно «смотреть» строго в камеру.

Поддерживает разработки по проблеме идентификации человека на основе термограммы лица фирма «Technology Recognition Systems» (США).

#### 2.4. ТЕХНОЛОГИЯ, ИСПОЛЬЗУЮЩАЯ РАСПОЗНАВАНИЕ ПО РАДУЖНОЙ ОБОЛОЧКЕ ГЛАЗА

Радужная оболочка глаза является уникальной для каждого человека биометрической характеристикой. Она формируется в период до полутора лет жизни человека и остается без изменений до конца его жизни, если здоровье человека не претерпело каких-либо принципиальных изменений. Последние отражаются на радужной оболочке в виде изменения ее цвета, появления на ней новых пигментных пятен, линий и кругов, изменений обвода оболочки, деформации зрачка и т. д. Эти особенности используются при иридоанализе для диагностики заболеваний или выявления предрасположенности к некоторым заболеваниям [Вел88, Вел98]. В биометрических системах идентификации человека используется полутонное изображение радужной оболочки, в котором некоторые пигментные изменения не так заметны.

**Методы идентификации человека по радужной оболочке глаза.** К настоящему моменту в системах идентификации человека по радужной оболочке сформировались два основных подхода, отличающиеся друг от друга способами представления образов радужки и соответствующих им образов-эталонов.

В первом подходе радужка (как новый образ или эталон) непосредственно выделяется из изображения глаза [Bor97, Wil97], а во втором — этим образом является матрица штрих-кодов, соответствующая структуре поверхности радужной оболочки [Sol97, Neg00].

В свою очередь, радужки в первом подходе могут быть представлены в двух формах:



- колец, выделенных из изображения глаза и определяющих только ту часть информационных знаков глаза, которая непосредственно относится к радужной оболочке [Wil97];

- прямоугольника, полученного из кольца радужной оболочки путем преобразования декартовой системы координат в полярную [Bor97, Kuc97a, Kuc98c].

Процесс получения кольца радужной оболочки непосредственно из изображения глаза показан на рис. 2.4.1.

Как видно из рис. 2.4.1, сначала на области зрачка определяется его центр (изображение 2), относительно которого определяются два радиуса:  $r$  и  $R$ . При этом  $r$  определяет радиус зрачка, а  $R$  — радиус до внешнего края радужной оболочки. Процесс определения значений  $r$  и  $R$  выполняется автоматически на основе пороговой обработки (сечений) и процедур вычисления контура. Полученные при этом кривые, определяющие контурные границы зрачка и границы радужки, не являются совсем точными и круглыми. Далее эти границы определяются более точно, таким образом определяется кольцо радужной оболочки. Заканчивает весь процесс процедура улучшения качества изображения кольца радужной оболочки (см. изображение 3).

Результат преобразования кольца радужной оболочки из декартовой системы координат в полярную форму представлен на рис. 2.4.2 [Bor97, Kuc97a, Kuc98c]. Весь процесс развертки радужной оболочки включает: определение местоположения, центра и контуров зрачка, определение величин радиусов окружности зрачка и окружности радужной оболочки, формирование координат полярной системы, и, наконец, преобразование каждого пикселя ра-

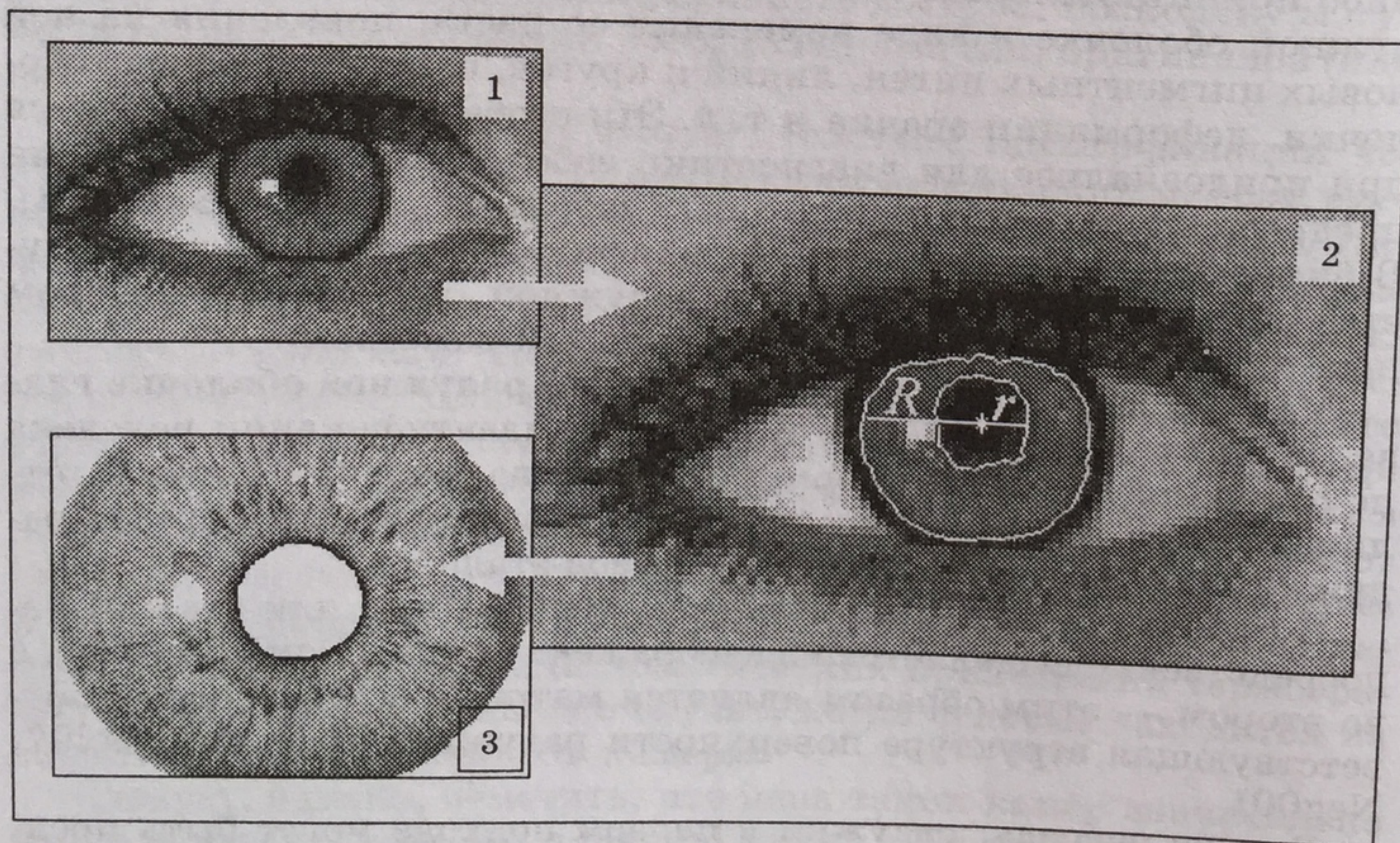


Рис. 2.4.1. Процесс получения кольца (отмечено цифрой 3) радужной оболочки



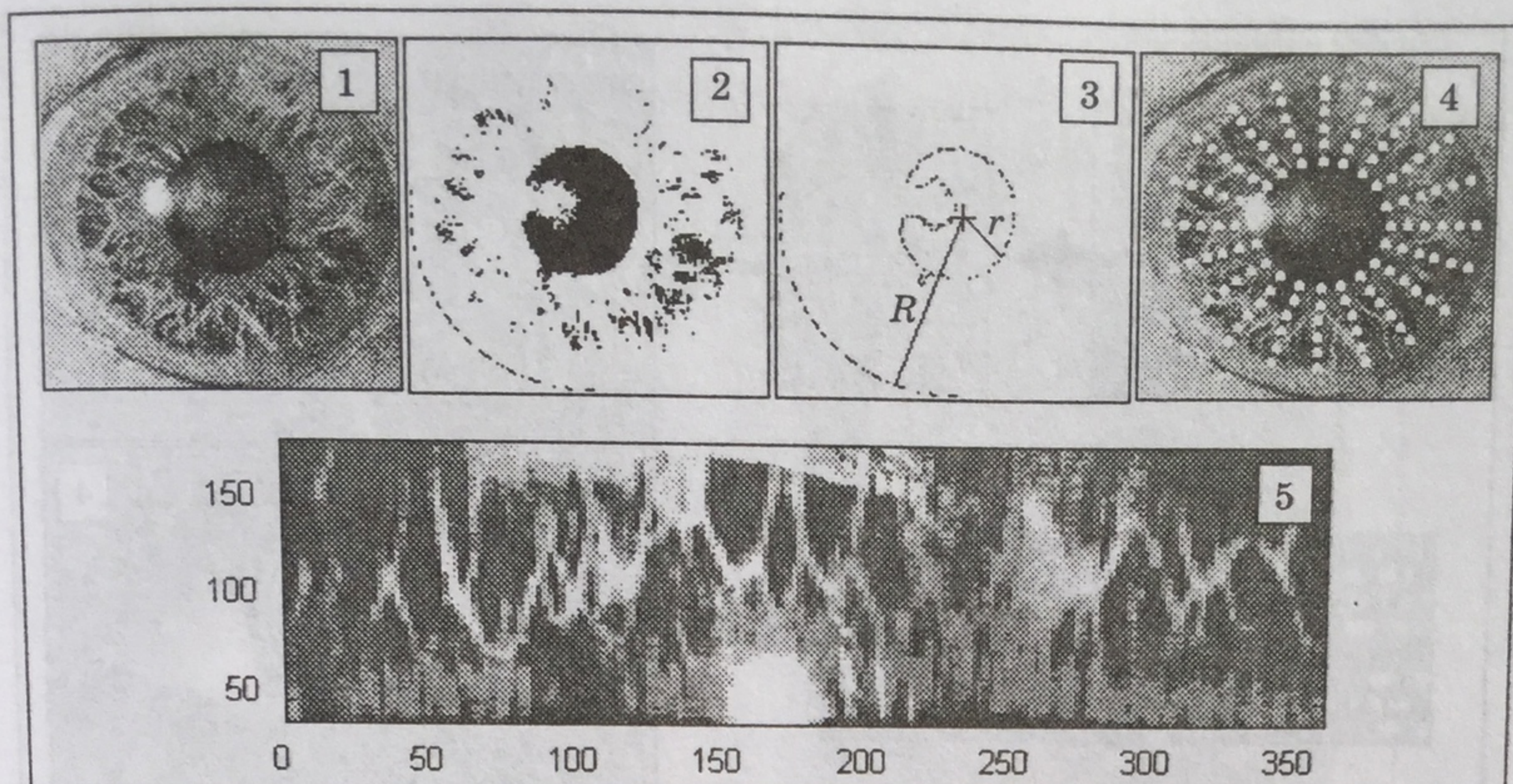


Рис. 2.4.2. Развертка радужной оболочки глаза в прямоугольную форму [Bor97]

радужной оболочки из его представления в декартовой системе координат в представление в полярной системе координат. Последовательность выполнения описанного процесса и получаемые при этом результаты отмечены на рисунке цифрами 1–5.

Поскольку целочисленные индексы декартовой системы координат не всегда строго соответствуют координатам полярной системы, на последнем этапе используется процедура интерполяции (нулевого порядка, двулинейная, или, например, билинейная [Kuc97a, Kuc98c]).

На рис. 2.4.2 в нижней его части, представляющей развертку радужки в прямоугольную форму, по оси  $X$  отложены углы полярной системы координат, а по оси  $Y$  — значения радиуса  $R$ .

Вторая форма представления радужных оболочек требует больших вычислительных затрат (в сравнении с затратами на представление радужки в форме кольца), однако является более удобной для практических применений.

Это обусловлено следующим:

- поворот изображения, преобразованного из декартовой системы координат в полярную, заменяется его циклическим сдвигом;
- используя логарифмическое представление координат в направлении радиуса  $R$  (вместо линейного), можно получить инвариант масштабного изменения радужной оболочки (в направлении радиуса  $R$ );
- используя два отмеченных выше свойства, получаем представление радужной оболочки глаза в форме инварианта поворота и масштаба, что значительно упрощает его сопоставление с аналогично представленным эталоном.

В качестве основного метода сравнения двух образов радужной оболочки в рассматриваемых случаях используется, как пра-



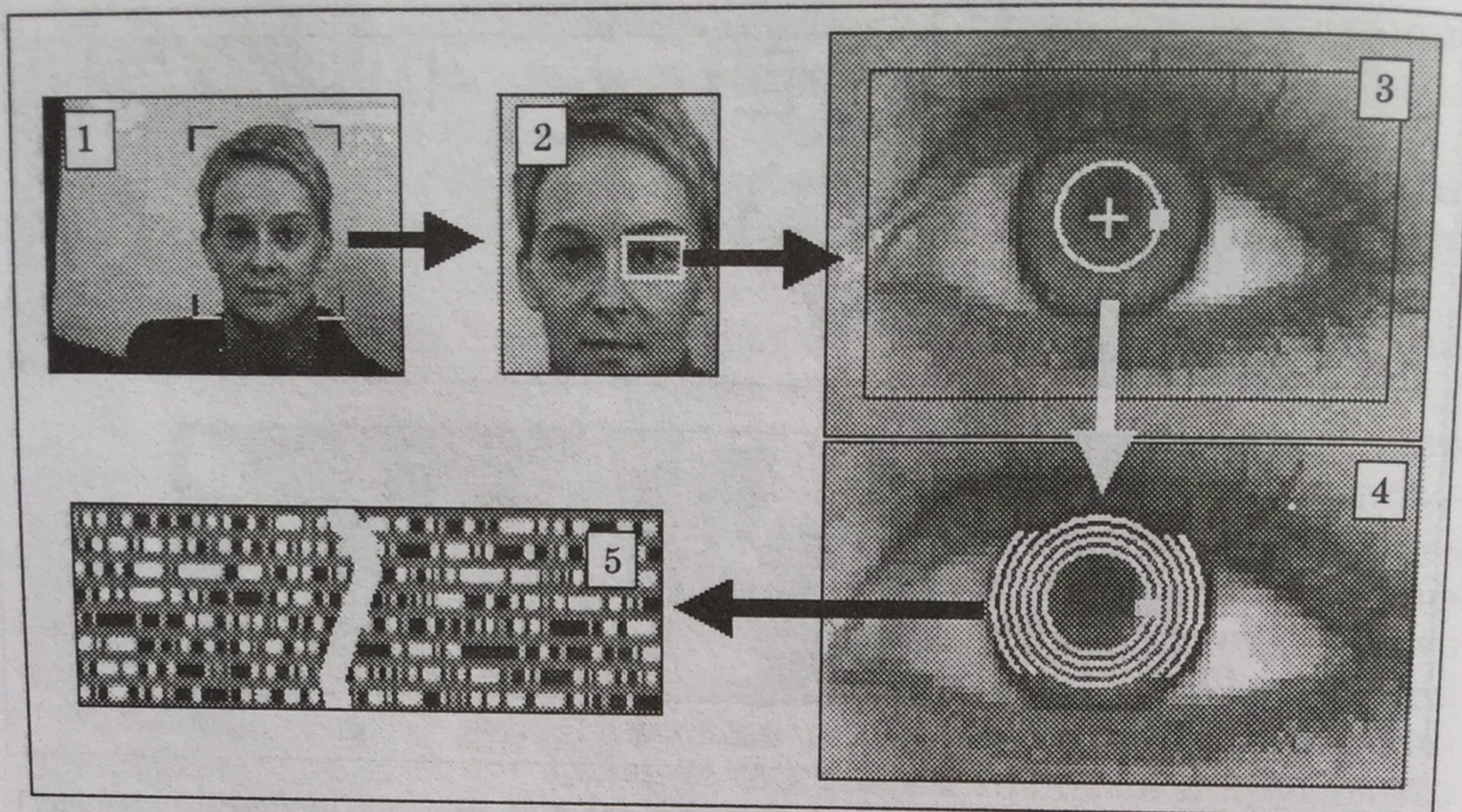


Рис. 2.4.3. Процесс получения матриц штрих-кодов радужной оболочки глаза

вило, взаимная корреляция. Некоторые результаты такого подхода представлены в [Bor97], где, в частности показано, что даже при значительных различиях в масштабе и повороте в изображениях двух одинаковых глаз коэффициент взаимной корреляции не был ниже 0,85.

Процесс получения матриц штрих-кодов, соответствующих структуре поверхности радужной оболочки, представлен схематически на рис. 2.4.3 [Sol97, Neg00]. Последовательность выполняемых при этом этапов и получаемых результатов преобразования исходных данных, отмечена цифрами от 1 до 5.

Этапы 1 – 3 связаны с выделением области лица на некоторой сцене, локализацией глаза на лице и зрачка на области глаза. На четвертом этапе выполняемого процесса на радужную оболочку глаза (нормированного по масштабу) накладывается фиксированная маска полярной системы координат. Значение каждого пикселя радужной оболочки в точке, соответствующей координатам маски, сравнивается с некоторым заданным порогом и в зависимости от результата сравнения записывается как «0» или «1» в определенное место матрицы штрих-кодов. Эта своеобразная бинарная матрица представлена на рис. 2.4.3 под номером 5.

Образ-эталон радужной оболочки (радужка зарегистрированного пользователя) записывается в память системы как файл размером 512 байтов. На этапе идентификации очередного клиента системы (или его аутентификации), сравнение каждой новой матрицы штрих-кодов с матрицей-эталоном может быть выполнено любым известным способом, хотя и здесь возможна взаимная корреляция. Другой способ сравнения (сопоставления) двух матриц — оценка расстояния между ними, например, в метрике  $L_0$  или  $L_2$ , хотя сдвиг на одну позицию (столбец или строку) одной матрицы относительно другой может привести к неправильному результату.



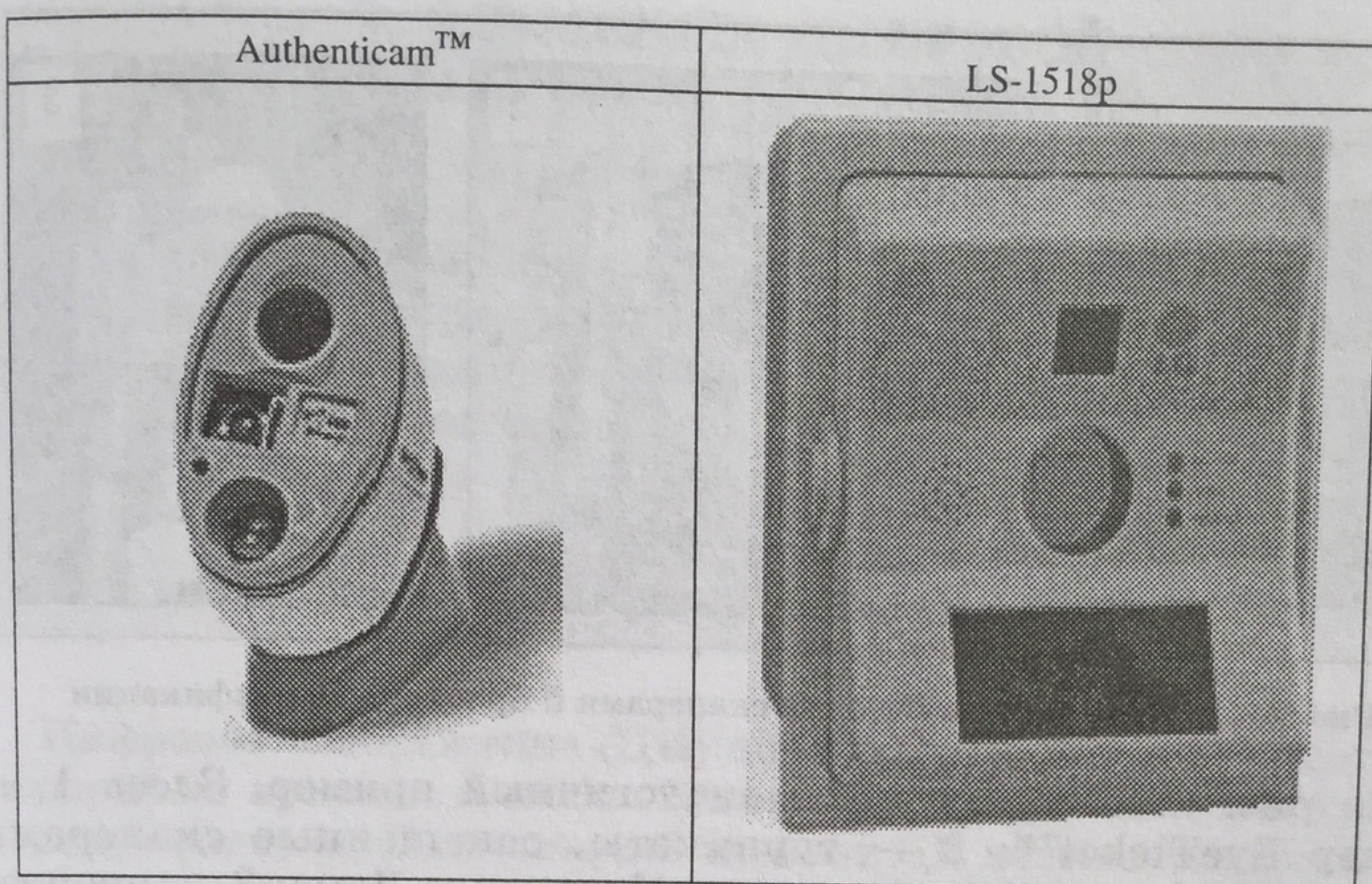


Рис. 2.4.4. Некоторые устройства для получения изображений глаз

**Устройства сканирования глаз.** На рис. 2.4.4 представлены некоторые устройства для получения изображений глаз и выделения из них образа радужной оболочки. Это автономные устройства, содержащие по две-три встроенные видеокамеры. Прибор «Authenticam™» фирмы «Iridian Technologies» предназначен в основном для индивидуального пользователя, в то время как прибор «LS-1518p» (совместное изделие компании «IriScan и LG») — для коллективного использования.

Прибор «Authenticam™» подключается к компьютеру через порт USB, имеет небольшие габариты, прост в обслуживании и может использоваться совместно с компьютерами Notebook.

Прибор «LS-1518p» используется в системе «IriScan's System 2100» идентификации человека, поставляемой компанией «IriScan» на рынок биометрических технологий. Как следует из проспекта компании «IriScan» на эту систему, база данных содержит 4000 эталонов. Регистрация пользователя занимает не более 30 секунд. Пользователь располагается напротив прибора «LS-1518p» на расстоянии 9 – 10 дюймов от него. Собственно процесс идентификации реализуется за 2 секунды. К системе «IriScan's System 2100» можно подключить более 256 удаленных терминалов с приборами «LS-1518p», в связи с чем эта система может использоваться для контроля доступа к распределенным информационным системам и базам данных. Более подробную информацию о системе можно найти на сайте [www.iriscan.com](http://www.iriscan.com).

Компания «EyeTicket Corporation» ([www.eyeticket.com](http://www.eyeticket.com)) привела на своем сайте пример использования собственных сканеров радужной оболочки глаза в системе контроля доступа EyeTicket™ на входе в Немецкий Дом в Сиднее.



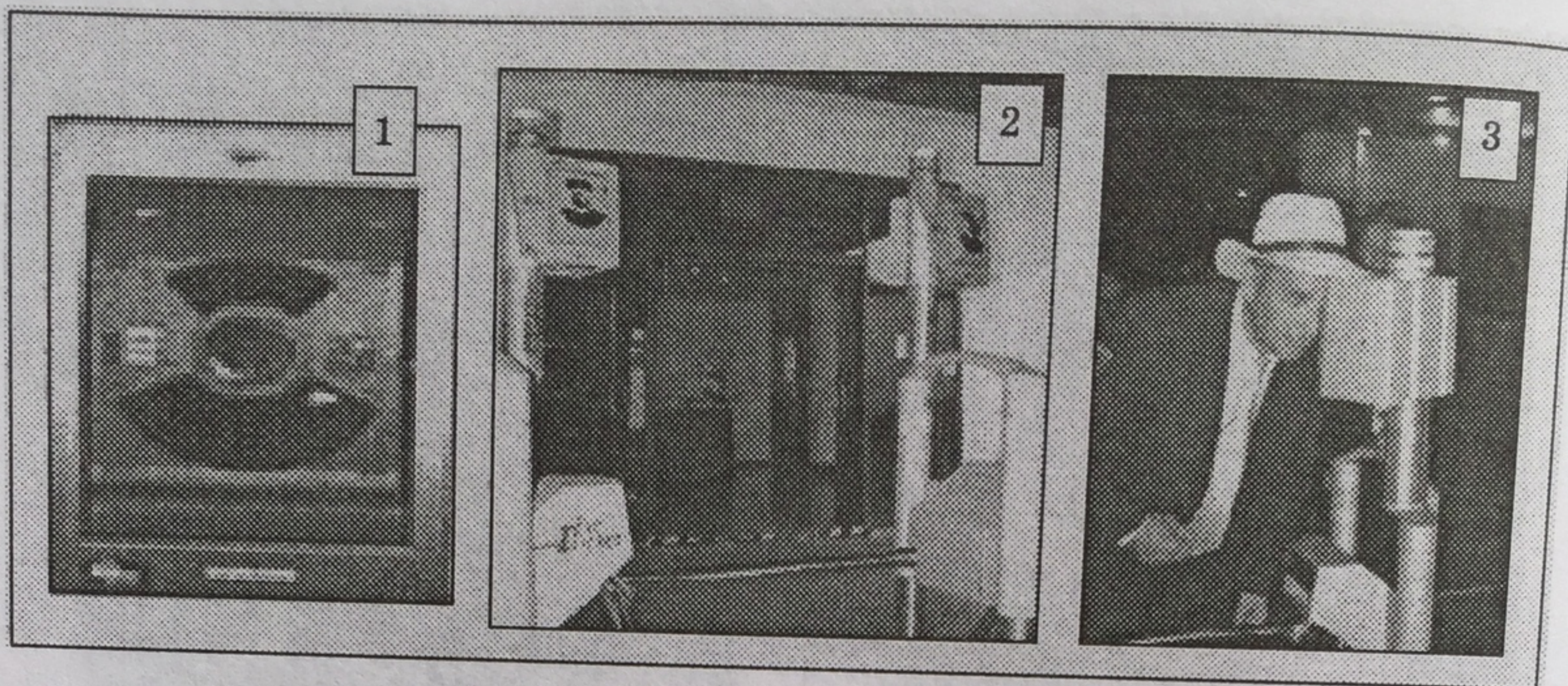


Рис. 2.4.5. Сканер, турникеты со сканерами и процесс аутентификации

На рис. 2.4.5 представлен аналогичный пример. Здесь 1 — сканер EyeTicket™; 2 — турникеты, оснащенные сканерами EyeTicket™, на входе в помещение Немецкого Дома; 3 — процесс аутентификации гостя Немецкого Дома.



### 3. ЭЛЕМЕНТЫ ЦИФРОВОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

В настоящей главе приведены краткие сведения о методах и алгоритмах, используемых при цифровой обработке изображений и в приложении к задачам распознавания образов, используемых в биометрических технологиях.

#### 3.1. ФОРМЫ И СПОСОБЫ ПРЕДСТАВЛЕНИЯ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Цифровое изображение (ЦИ) появляется в результате дискретизации по времени и квантования по уровню аналогового двумерного сигнала. Наименьший элемент ЦИ называется пикселем (pixel). Цифровое изображение представляется в общем случае прямоугольной таблицей, имеющей  $N$  строк и  $M$  столбцов, где каждый элемент есть пиксель. Эту таблицу можно также записать в виде матрицы размером  $N \times M$  элементов.

Для графического представления координат пикселей ЦИ используется несколько различных способов.

В задачах, связанных с распознаванием образов, одно и то же цифровое изображение может быть представлено различными способами как в декартовой, так и в полярной системах координат.

На рис. 3.1.1 приведены два способа представления цифрового изображения в декартовой системе координат.

Левая координатная система соответствует оси  $X$ , идущей слева направо, а правая координатная система — оси  $Y$ , идущей снизу вверх. Поэтому в левом нижнем углу матрицы, представляющей ЦИ в левой координатной системе, находится пиксель с координатами  $(1, 1)$ , а в правом верхнем углу — пиксель с координатами  $(N, M)$ .

В правой координатной системе порядковое перечисление пикселей ЦИ начинается с верхнего левого угла соответствующей матрицы и заканчивается в правом нижнем углу. И хотя такое представление координат отличается от общепринятого представления в двумерной левосторонней декартовой системе, оно наиболее часто используется в отображении цифровых изображений на плоскости  $XY$ .

Расстояние  $d$  между двумя пикселями, имеющими координаты  $(x_1, y_1)$  и  $(x_2, y_2)$ , определяется следующим образом:

$$d_{1,2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (3.1.1)$$

Пусть нам задана матрица 8-го порядка,



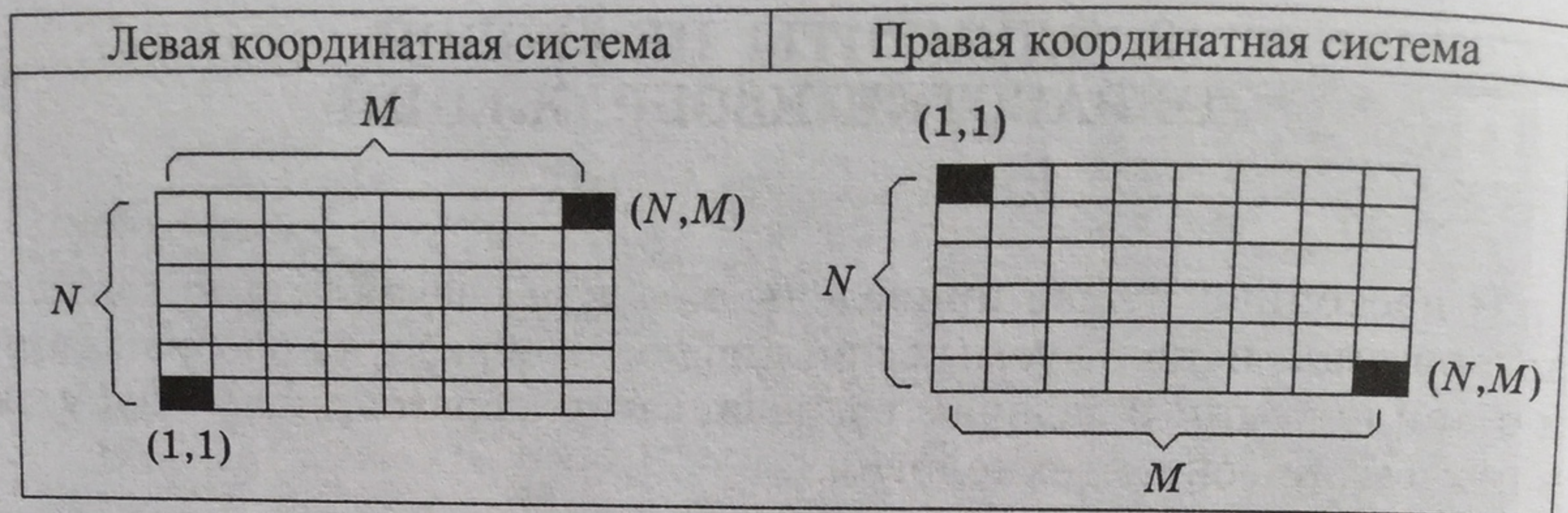


Рис. 3.1.1. Два способа представления цифрового изображения в декартовой системе координат

$$\begin{bmatrix}
 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
 2 & 4 & 6 & 8 & 10 & 12 & 14 & 16 \\
 3 & 6 & 9 & 12 & 15 & 18 & 21 & 24 \\
 4 & 8 & 12 & 16 & 20 & 24 & 28 & 32 \\
 5 & 10 & 15 & 20 & 25 & 30 & 35 & 40 \\
 6 & 12 & 18 & 24 & 30 & 36 & 42 & 48 \\
 7 & 14 & 21 & 28 & 35 & 42 & 49 & 56 \\
 8 & 16 & 24 & 32 & 40 & 48 & 56 & 64
 \end{bmatrix}, \quad (3.1.2)$$

определяющая изображение размером  $8 \times 8$  пикселей.

В графической форме это же изображение в декартовой системе координат представлено на рис. 3.1.2. Здесь буквой а определено изображение (3.1.2) в левосторонней системе координат, а буквой b — то же изображение в правосторонней системе координат.

Матрица (3.1.2) может быть графически представлена и в трехмерной декартовой системе координат. В этом случае элементы матрицы размещаются на плоскости XY, а значения этих элемен-

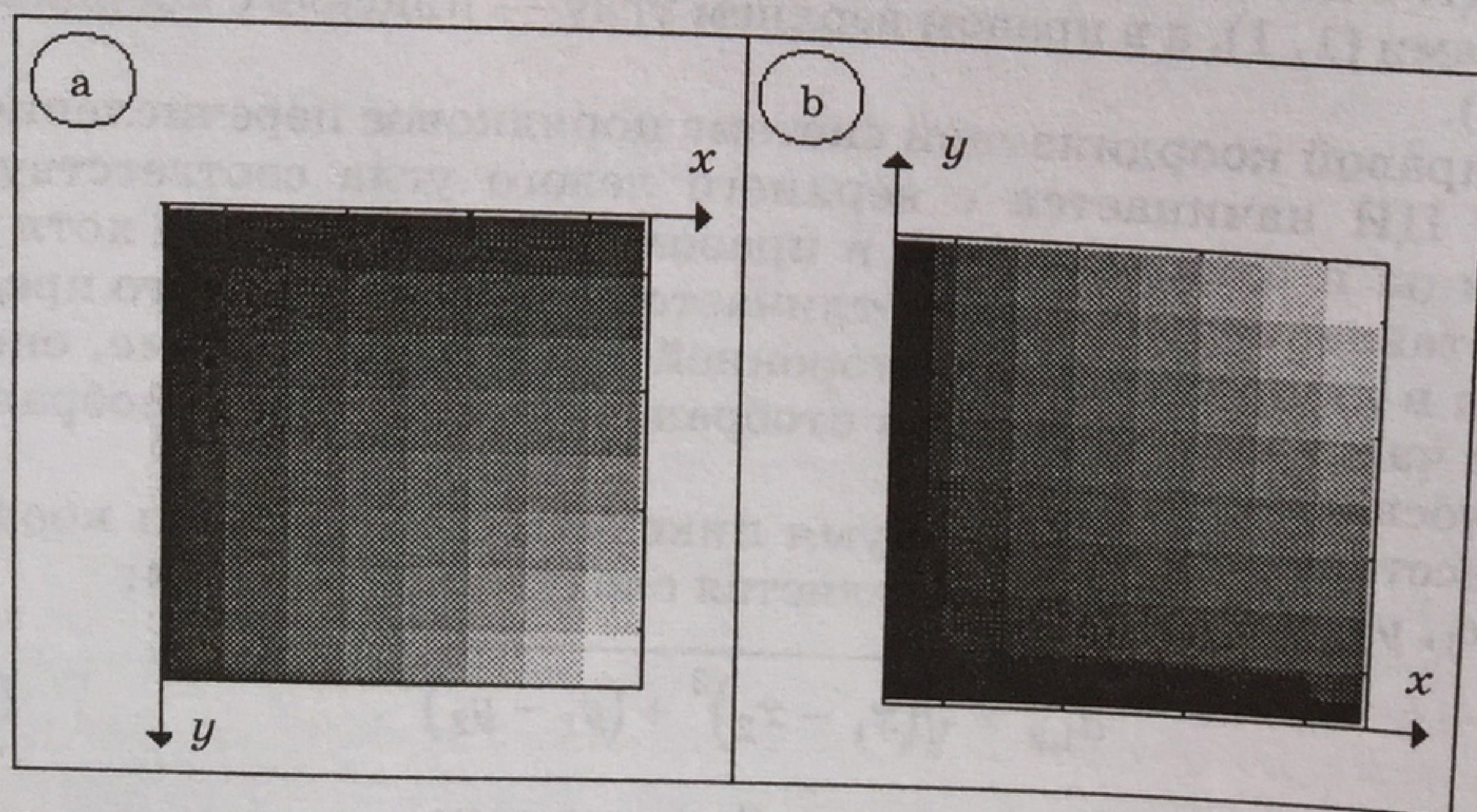


Рис. 3.1.2. Матрица (3.2), представленная как цифровое изображение



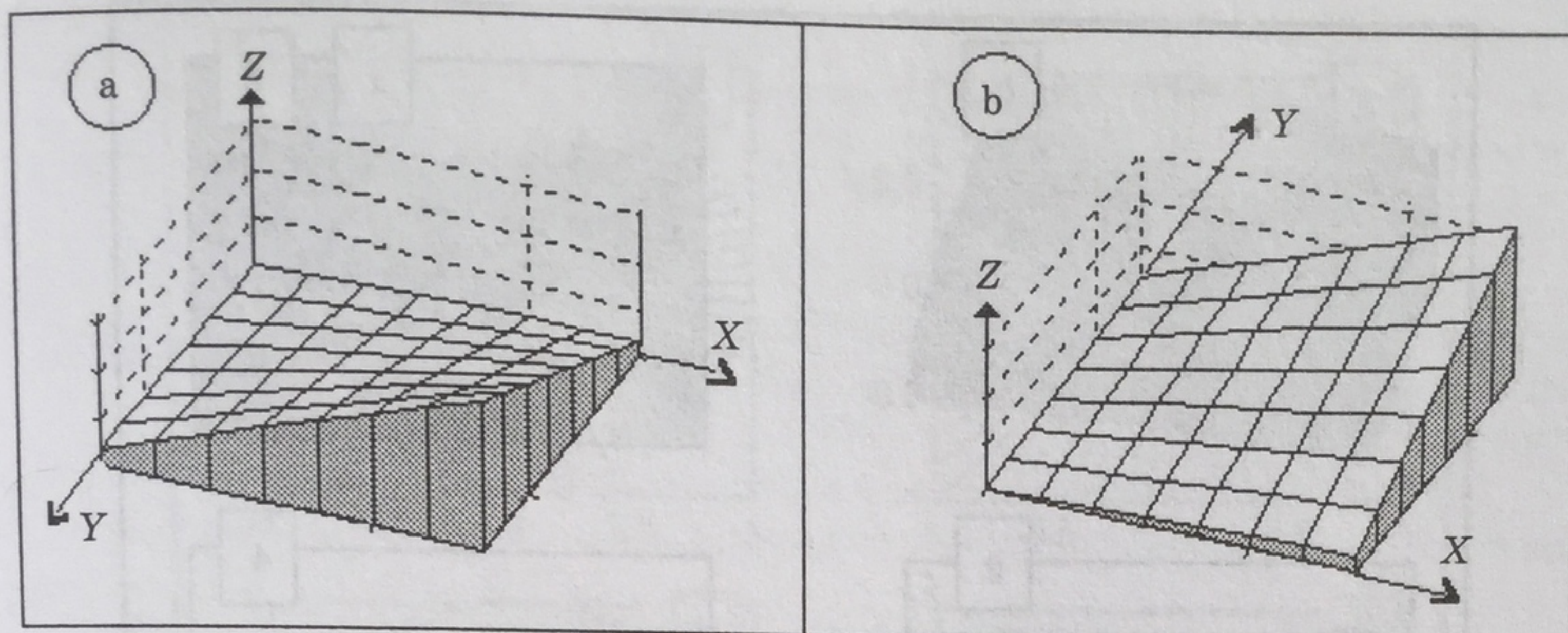


Рис. 3.1.3. Матрица (3.2), представленная в форме 3D образа

тов откладываются по оси  $Z$ . Результат такого представления показывает рис. 3.1.3.

На рис. 3.1.3 буквой а определено изображение (3.1.2) в левосторонней трехмерной декартовой системе координат, а буквой б — то же изображение в правосторонней трехмерной декартовой системе координат.

По терминологии [Пав86] изображения, представленные в форме 3D, относятся к цифровым изображениям класса «0». Чтобы не путать понятие «класс из [Пав86]» с общепринятым понятием класса, используемым в распознавании образов, заменим первое термином с п о с о б. И теперь отметим, что в [Пав86] введено пять способов описания и представления цифровых изображений, из которых способы 1 – 4 предназначены для представления изображений в форме 2D. Последние разделены на: полутоновые бинарные, контурные и изображения, представленные отдельными точками или своими локальными областями.

Пример графического представления одного и того же изображения (участка лица) в форме 3D и 2D приведен на рис. 3.1.4. Форма 2D представлена полутоновым и контурным изображениями, а также изображением с локальными областями лица. Над каждым из изображений указан способ его представления (класс по [Пав86], которому принадлежит изображение).

Благодаря введенной в [Пав86] классификации можно описывать соотношения между изображениями на входе и выходе систем обработки изображений. Так, например, вертикальная проекция 3D-изображения на плоскость  $XY$  позволяет получить изображение в форме 2D, а срезы последнего при различных значениях порогов переводят полутоновое 2D-изображение в бинарное. В свою очередь, бинарное изображение легко преобразуется в контур при помощи какой-либо процедуры контуризации и т. д. Такое нисходящее преобразование ЦИ часто используется в задачах распознавания образов при выделении признаков на изображениях.

Выбор другой (не декартовой) системы координат для графического представления ЦИ определяется конкретной областью при-



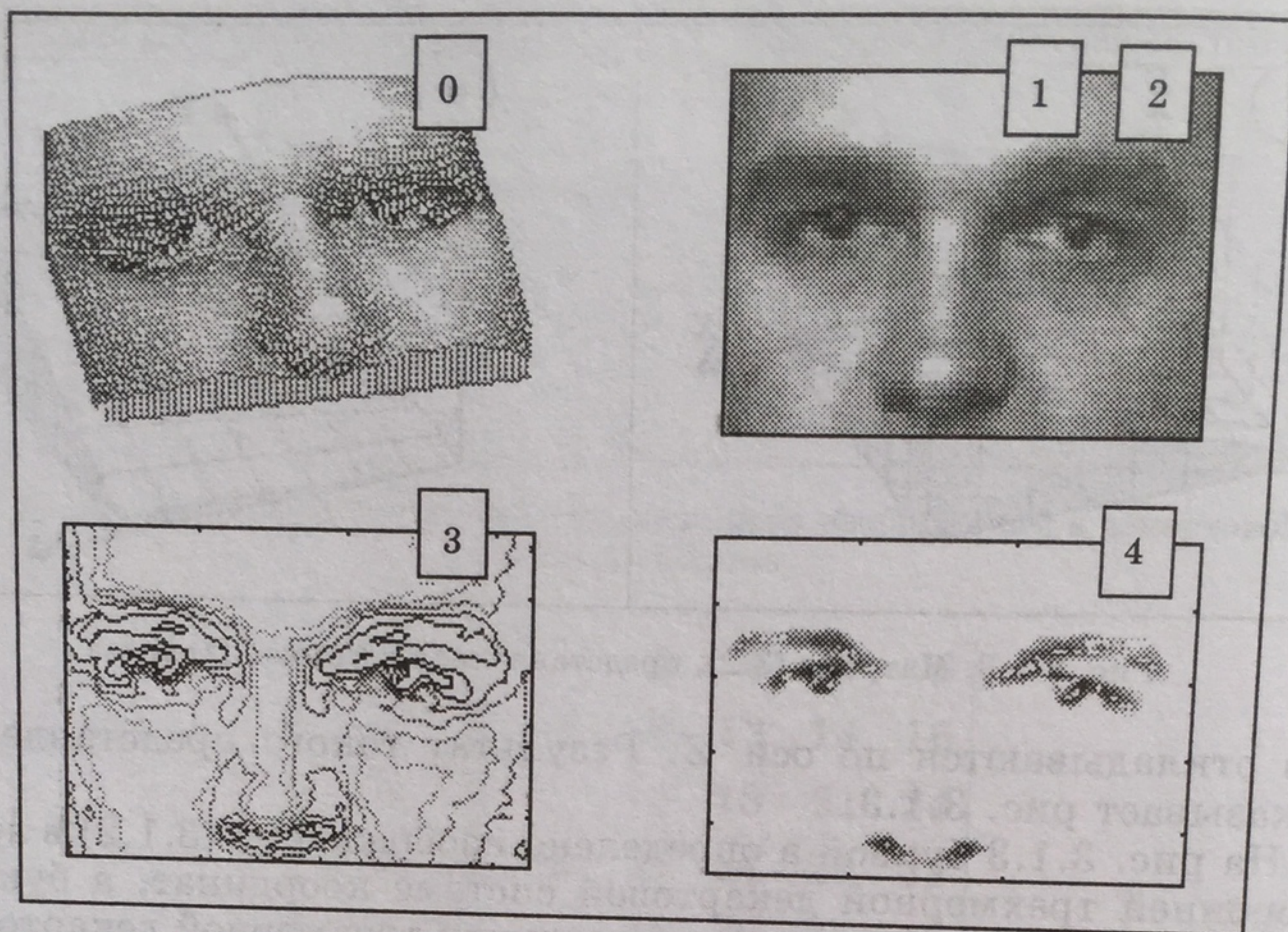


Рис. 3.1.4. Пять способов представления одного и того же изображения

ложений и вытекающей из нее спецификой решаемой задачи. Например, в преобразовании Фурье—Меллина, широко используемом при распознавании образов [Ben98, Che94], трансформируется спектр цифрового изображения из декартовой системы координат в полярную «прямоугольную» систему координат. Последняя является как бы разверткой обычной полярной системы координат. Особенности такого представления ЦИ показаны на рис. 3.1.5.

Координаты пикселей из одной системы в другую пересчитываются по следующим соотношениям:

$$\rho = \sqrt{x^2 + y^2}; \operatorname{tg}(\varphi) = \frac{y}{x}, x \neq 0; x = \rho \cos(\varphi); y = \rho \sin(\varphi). \quad (3.1.3)$$

От способа дискретизации полярного радиуса (равномерный или неравномерный) зависит способ размещения координат на

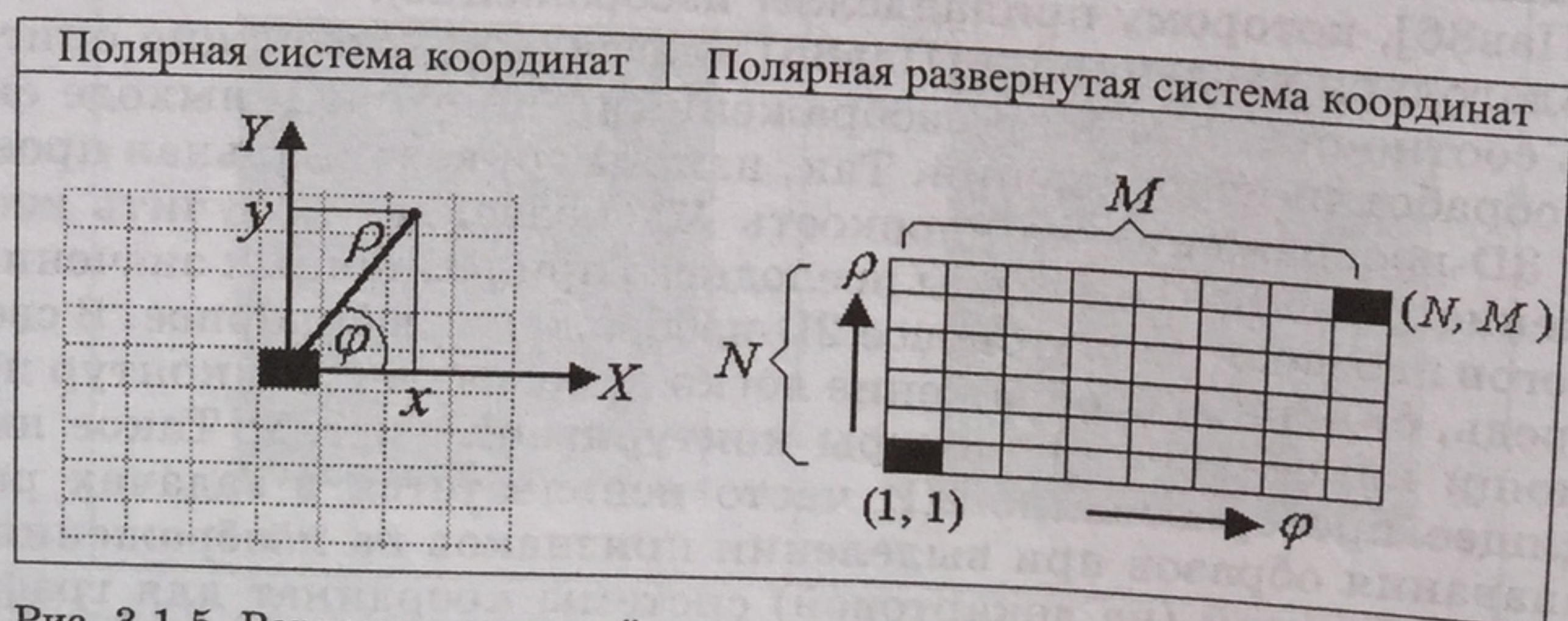


Рис. 3.1.5. Варианты полярной системы координат и размещение пикселей



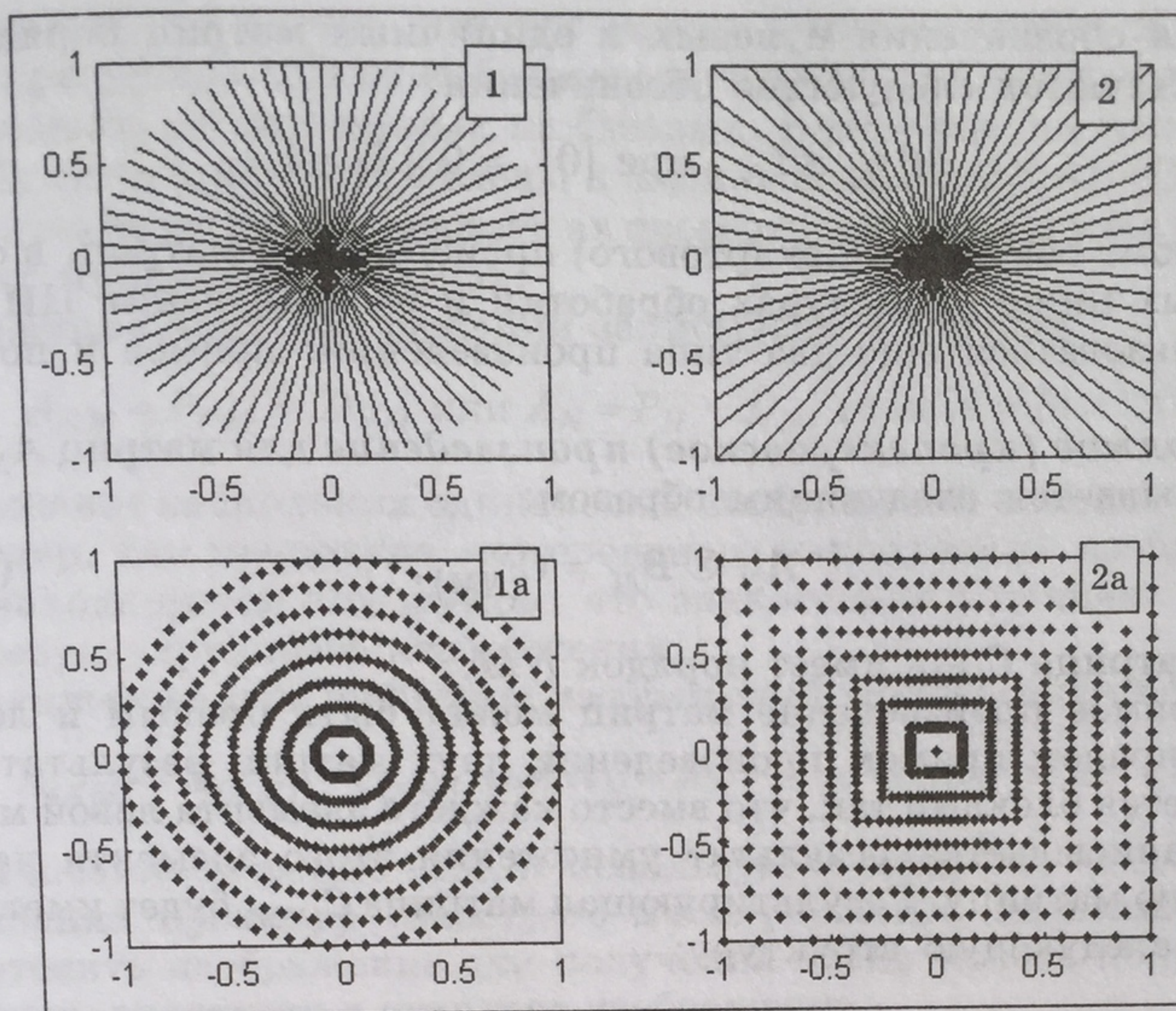


Рис. 3.1.6. Способы размещения полярных координат на цифровом изображении

поле ЦИ. Оба эти способа, широко применяемые в практике цифровой обработки изображений, показаны на рис. 3.1.6.

### 3.2. СИСТЕМА ОБОЗНАЧЕНИЙ, ИСПОЛЬЗУЕМАЯ В АЛГОРИТМАХ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Пусть некоторое цифровое изображение записано матрицей размером  $N \times M$  пикселей. Тогда каждый пиксель этого изображения запишем в виде [Прэ82, Даг83, Кис98а, Гол99]

$$p(n, m), \quad (3.2.1)$$

где  $n = 1, 2, \dots, N$  и  $m = 1, 2, \dots, M$ .

Все изображение можно записать в форме  $P_{N \times M}$  или в форме  $P_N$  (для случая  $N = M$ ). При этом нижний индекс (выражение) в обозначении матрицы всегда будет обозначать ее порядок (или размер для случая  $N \neq M$ ).

К каждой матрице, определяющей ЦИ, можно применить операции транспонирования, обращения, комплексного сопряжения, возведения в степень и т. д., которые записываются наиболее принятыми для этих операций обозначениями. Например:

$$[P_N]^T, [P_N]^{-1}, [P_N]^*, [P_N]^k. \quad (3.2.2)$$



Для обозначения нулевых и единичных матриц порядков  $N$  используются следующие обозначения:

$$[0]_N \text{ и } I_N, \text{ при } [0]_1 = 0 \text{ и } I_1 = 1. \quad (3.2.3)$$

Кроме обычного (декартового) произведения матриц, в обсуждаемых ниже процедурах обработки и распознавания ЦИ будут использоваться еще два типа произведений: прямое и поточечное.

**Прямое (кронекеровское) произведение** для матриц  $A_N$  и  $B_M$  записывается следующим образом:

$$A_N \otimes B_M = C_{(NM)}, \quad (3.2.4)$$

где матрица  $C_{(NM)}$  имеет порядок  $NM$ .

Прямое произведение матриц может быть правым и левым. При правом прямом произведении двух матриц результат формируется блоками так, что вместо каждого элемента левой матрицы записывается результат умножения этого элемента на всю правую матрицу. Результирующая матрица  $C_{(NM)}$  будет иметь при этом следующую структуру:

$$C_{(NM)} = \begin{bmatrix} a_{11}B_M & \cdot & \cdot & a_{1N}B_M \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{N1}B_M & \cdot & \cdot & a_{NN}B_M \end{bmatrix}. \quad (3.2.5)$$

**Поточечное произведение** матриц  $A_N$  и  $B_N$  записывается в форме

$$A_N \odot B_N = C_N, \quad (3.2.6)$$

где матрица  $C_N$  порядка  $N$  определяется следующим образом:

$$C_N = \begin{bmatrix} a_{11}b_{11} & \cdot & \cdot & a_{1N}b_{1N} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{N1}b_{N1} & \cdot & \cdot & a_{NN}b_{NN} \end{bmatrix}. \quad (3.2.7)$$

### 3.3. БАЗОВЫЕ ПРОЦЕДУРЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Представим некоторые простые (базовые) процедуры, используемые в задачах обработки цифровых изображений, в том числе процедуры сложения, вычитания и умножения цифровых изображений, процедуры, основанные на логических операциях (операциях «или» и сложения по модулю 2). При этом нашей целью здесь является представление процедур обработки изображений в век-



торно-матричной форме, которую можно непосредственно использовать при реализации этих базовых процедур, а также других более сложных, но основанных на базовых. Некоторые из этих процедур будут также представлены в языке пакета MATLAB [Пот99, Руд00], что позволит реализовать их проверку и использование в конкретных прикладных задачах.

**Сложение** двух изображений записывается в форме

$$A_{NM} = P_{NM} + D_{NM} \text{ или } A_N = P_N + D_N, \text{ если } N = M. \quad (3.3.1)$$

Сложение нескольких одинаковых изображений используется, например, как процедура «когерентного накопления» изображений, находящихся под шумом, что значительно улучшает качество результирующего изображения.

**Вычитание** двух цифровых изображений записывается в форме

$$A_{(NM)} = X_{(NM)} - Y_{(NM)} \text{ или } A_N = X_N - Y_N, \text{ если } N = M. \quad (3.3.2)$$

Вычитание в форме (3.3.2) используется довольно часто для реализации процедур «неострого маскирования», позволяющих подготовить изображение для получения очень четких контуров объектов, входящих в исходное изображение.

**Умножение** поточечное цифровых изображений можно реализовать следующим образом:

$$Y_N = X_N \odot X_N \odot \dots \odot X_N, \quad (3.3.3)$$

что обычно используется для улучшения качества изображений.

Заметим, что при выполнении операций (3.3.1) – (3.3.3) необходимо следить за значениями (яркости) пикселей, величины которых должны находиться в пределах заданного диапазона. Например, для полутоновых изображений этот диапазон находится в пределах от 1 до 255.

Использование логических операций покажем на примере выделения контура бинарного изображения. Итак, пусть матрица  $X_N$  порядка  $N$ , представляет бинарное изображение. Два примера таких изображений приведены на рис. 3.3.1 под номерами 1 и 3. Вычисление контура реализуем по следующей процедуре [Кис98а]:

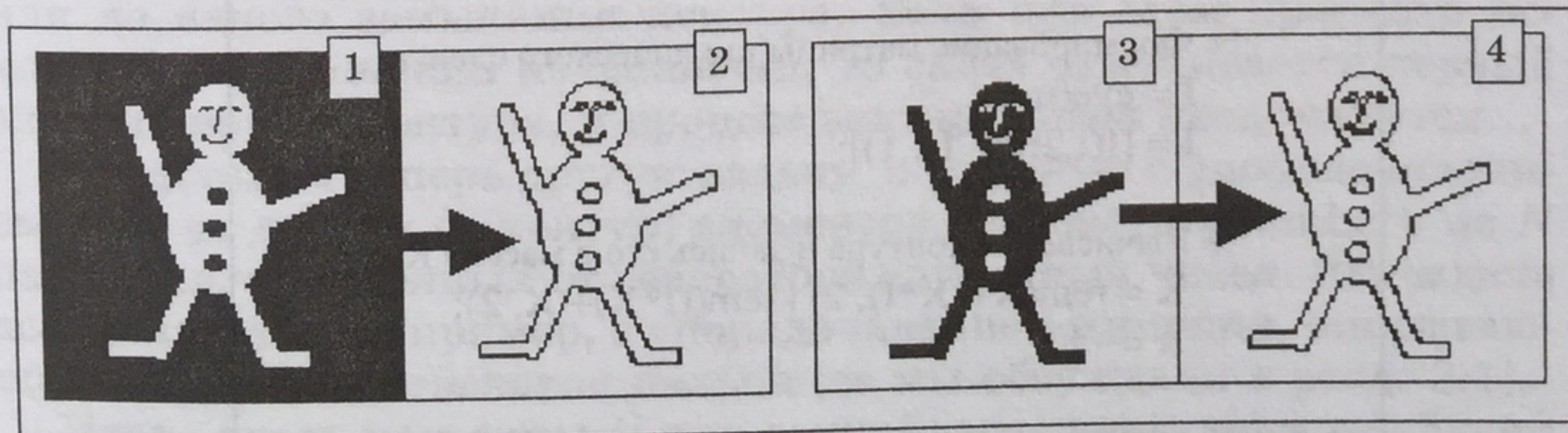


Рис. 3.3.1. Результаты вычисления контуров по процедуре (3.3.4)



$$K_N = \left( X_N^{(\Leftarrow)} + X_N \right)_{\text{mod } 2} \vee \left( X_N^{(\Downarrow)} + X_N \right)_{\text{mod } 2}, \quad (3.3.4)$$

где  $X_N^{(\Leftarrow)}$  — матрица  $X_N$  с циклически сдвинутыми влево (или вправо) на одну позицию столбцами;  $X_N^{(\Downarrow)}$  — матрица  $X_N$  с циклически сдвинутыми вверх (или в низ) на одну позицию строками;  $K_N$  — матрица порядка  $N$ , определяющая вычисляемый контур;  $\vee, \text{mod } 2$  — знаки операции «или» и сложения по модулю 2.

Результат вычисления контуров для двух приведенных на рис. 3.3.1 изображений представляют два новых изображения — 2 и 4 соответственно. Из результатов видно, что независимо от фона исходных изображений (образ на черном или белом фоне) процедура приводит к одному и тому же результату.

Программа вычисления контура бинарного изображения по процедуре (3.3.4), написанная в языке пакета MATLAB, приведена ниже.

Заметим, что выбор любой пары сдвигов из возможных вариантов (влево + вверх, влево + вниз, вправо + вверх и, наконец, вправо + вниз) приводит к сдвигу границ вычисляемых контуров на одну позицию, что является ошибкой метода (3.3.4). Однако в рассматриваемых здесь приложениях эта методическая ошибка не столь принципиальна. Результат, представленный на рис. 3.3.1, получен на основе приведенной программы 1.

**Векторизация.** В разд. 2.3 мы отмечали возможность использования для распознавания лиц эластичных моделей форм лица

Программа 1

```
function K = kontur(X)

% Вычисление контура бинарного изображения,
% представленного квадратной матрицей X

% Проверка размера матрицы X
[m,n] = size(X);
if m ~= n return; end;

% Формирование матрицы циклического сдвига
I = eye(m);
I = [I(:,2:m) I(:,1)];

% Вычисление контура и запись его в массив K
K = rem(X+(X*I), 2) | rem((I*X)+X, 2);
K = ~ K;

% Конец процедуры
```



(ЭМФЛ). ЭМФЛ описывает контуры основных деталей лица (овал лица, контуры носа, глаз и бровей, контуры губ) в форме вектора, содержащего координаты точек контура. При автоматическом построении ЭМФЛ неизбежен переход от контура (как бинарного образа) к его векторному представлению. Это и есть операция векторизации контура.

Подходов и методов, предназначенных для реализации операций векторизации контуров, известно достаточно много (см., например, [Пав86, Прэ82, Янш94]). Остановимся лишь на одном, который можно условно назвать «методом поиска ближайшего соседа» и который предназначен для случаев, когда бинарное изображение состоит из нескольких контуров. Подробно этот метод изложен и проиллюстрирован в [Кис98а].

Суть метода состоит в следующем. Положим, что пиксели, соответствующие элементам контура, имеют значение, равное 1. Пиксели фона при этом имеют значения, равные 0. Находим первый «ненулевой» пиксель путем, например, сканирования по строкам исходного изображения. Отметим координаты  $(x_n, y_n)$  этого пикселя как начало первого контура. Далее определим эти координаты как центр маски размером  $3 \times 3$  пикселя и будем просматривать все оставшееся пространство в поле маски. Фактически среди восьми ближайших к центру маски пикселей отыскиваем такие, которые имеют ненулевые значения. В случае обнаружения такового, определяем расстояние до него, используя соотношение (3.1.1). По окончании всего просмотра в поле маски следующим элементом контура маски будем считать тот, расстояние до которого окажется минимальным. Это и есть «ближайший сосед». Координаты центра маски переписываем в результирующий вектор, центр маски передвигаем на «найденного соседа» и весь процесс поиска нового «ближайшего соседа» повторяется. Найденные элементы контура из дальнейшего рассмотрения исключаются, например, «обнулением» соответствующих пикселей. Процесс векторизации текущего (локального) контура заканчивается тогда, когда последний найденный элемент контура будет соседом пикселя, определенного как начало контура. Если в поле рамки  $3 \times 3$  пикселя не обнаружено ни одного «ненулевого» соседа, сторона рамки увеличивается до 5 (7 или 9 и т. д.) пикселей, вплоть до некоторого порога увеличения или до явного замыкания контура. Если при этом просмотр исходного изображения не закончен, то снова отыскивается первый элемент нового контура, и процесс векторизации продолжается...

Представим теперь другую задачу, связанную с упорядочиванием друг за другом (в контур) элементов вектора, состоящего из  $N$  пар чисел, полученных от генератора случайных чисел. Эта задача ассоциируется, например, с упорядочиванием минут, описывающих некоторый отпечаток пальца (ее мы обсуждали в разд. 2.1).

Итак, пусть нам даны  $N$  пар случайных чисел, которые будем рассматривать как координаты  $x$  и  $y$  на некоторой плоскости  $XY$ .



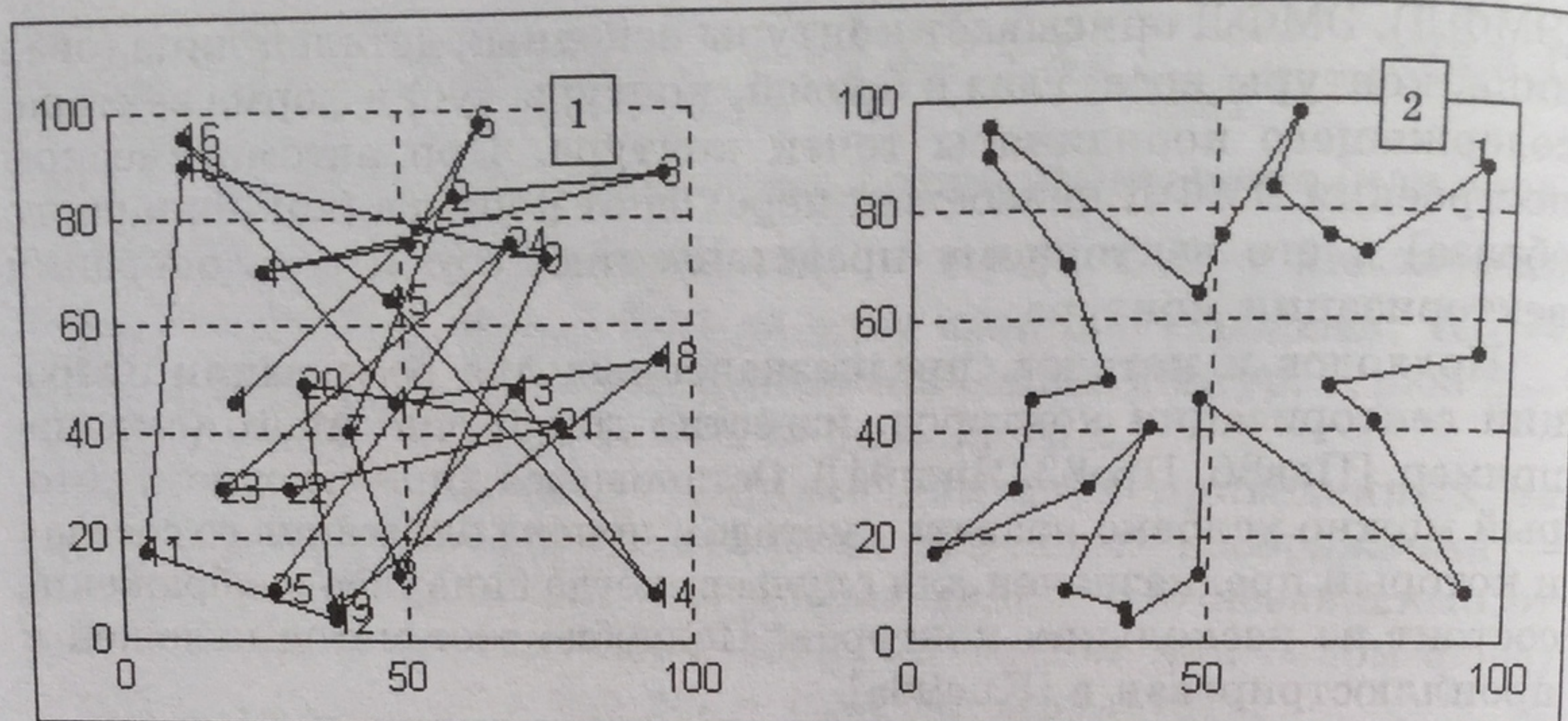


Рис. 3.3.2. Получение контура из случайно сгенерированных пар координат

Это, например, могут быть координаты всех пикселей (отдельного контура), значения которых не равны 0. Если соединить между собой эти координаты (на плоскости  $XY$ ) в той последовательности, в которой они сгенерированы (или были заданы), то получится фигура, подобная той, что показана на рис. 3.3.2 под номером 1. Нам же нужно получить фигуру, представленную на рис. 3.3.2 под номером 2.

Программа в языке пакета MATLAB, реализующая представленную задачу, приведена ниже (Программа 2).

Запишем все исходные координаты в форме вектора  $X$ , состоящего из  $N$  комплексных чисел:

$$X = [(x_1 + jy_1), (x_2 + jy_2), \dots, (x_N + jy_N)]. \quad (3.3.5)$$

С учетом (3.3.5) в программе 2 реализованы следующие действия.

1. Исходные данные центрированы относительно среднего значения:

$$X_0 = X - \bar{X}, \quad (3.3.6)$$

где  $\bar{X}$  — среднее значение вектора  $X$ .

2. Вычислены полярные углы для всех пар центрированных координат:

$$\theta = \arctg(\text{imag}(X_0)/\text{real}(X_0)). \quad (3.3.7)$$

3. Выполнена сортировка полярных углов по возрастанию и запоминается порядок их пересортировки (например, в векторе  $\text{index}$ ).

4. Исходные координаты сортируются в соответствии с возрастанием полярного угла:

$$W = X(\text{index}). \quad (3.3.8)$$



```

function W=wektor(X)

% Упорядочивание друг за другом (в контур) ...
% элементов вектора.
% X – вектор неупорядоченных комплексных чисел
% W – вектор упорядоченных друг за другом
% комплексных чисел

% Центрирование исходных данных
jo t= sqrt(-1); x = real(X); y = imag(X);
len = length(x);
xcentr = fix(sum(x)/len); ycentr = fix(sum(y)/len);
x = x-xcentr; y = y-ycentr;

% Вычисление полярных координат
[tet,ro] = cart2pol(x,y);
% Сортировка по полярному углу
[a,index] = sort(tet);
X1 = x(index); Y1 = y(index);
% Формирование результата
W = X1+xcentr + jot*(Y1+ycentr);
% Замыкание контура
W = [W W(1)];

% Конец процедуры

```

**Развертка контурного изображения.** В задачах распознавания изображений (например, рукописных знаков, букв, цифр и др.) довольно часто используются инварианты аффинных преобразований контурных объектов на плоскости  $XY$ : поворота, масштаба, сдвига [Rau92, Кис99а]. Представим здесь UNL — инварианты [Rau92], получаемые в результате преобразования контурных объектов из декартовой системы координат в полярную, и их развертки в прямоугольную форму, показанную на рис. 3.1.5 с правой стороны.

1. Запишем исходные координаты контура в форме вектора  $X$ , состоящего из  $N$  комплексных чисел:

$$X = [(x_1 + jy_1), (x_2 + jy_2), \dots, (x_N + jy_N)]. \quad (3.3.9)$$

2. Центрируем исходные данные относительно среднего значения:

$$X_0 = X - \bar{X}, \quad (3.3.10)$$

где  $\bar{X}$  — среднее значение вектора  $X$ .



```

function [R,TETA] = unl(X)

% Вычисление UNL- инвариантов
% X      - вектор неупорядоченных комплексных чисел (координат);
% R      - полярный радиус;
% TETA   - полярный угол (в градусах)

% Центрирование исходных данных
Mx = mean(X);
Xo = X-Mx;

% Вычисление полярных координат
[tet,ro] = cart2pol(real(Xo), imag(Xo));

% Сортировка по полярному углу, перевод в градусы
[TETA,index] = sort(tet);
TETA = (TETA-TETA(1))*360/(2*pi);

% Нормировка радиуса по масштабу к "1"
Mmax = max(abs(Xo));
R = ro(index)/Mmax;

% Конец процедуры

```

3. Вычислим нормированный полярный радиус и полярные углы для всех пар центрированных координат:

$$\left. \begin{aligned} R &= \text{abs}(X_0) / \max(X_0); \\ \Theta &= \text{arctg}(\text{imag}(X_0) / \text{real}(X_0)), \end{aligned} \right\} \quad (3.3.11)$$

где  $\max(X_0)$  — максимальный элемент в векторе  $X_0$ .

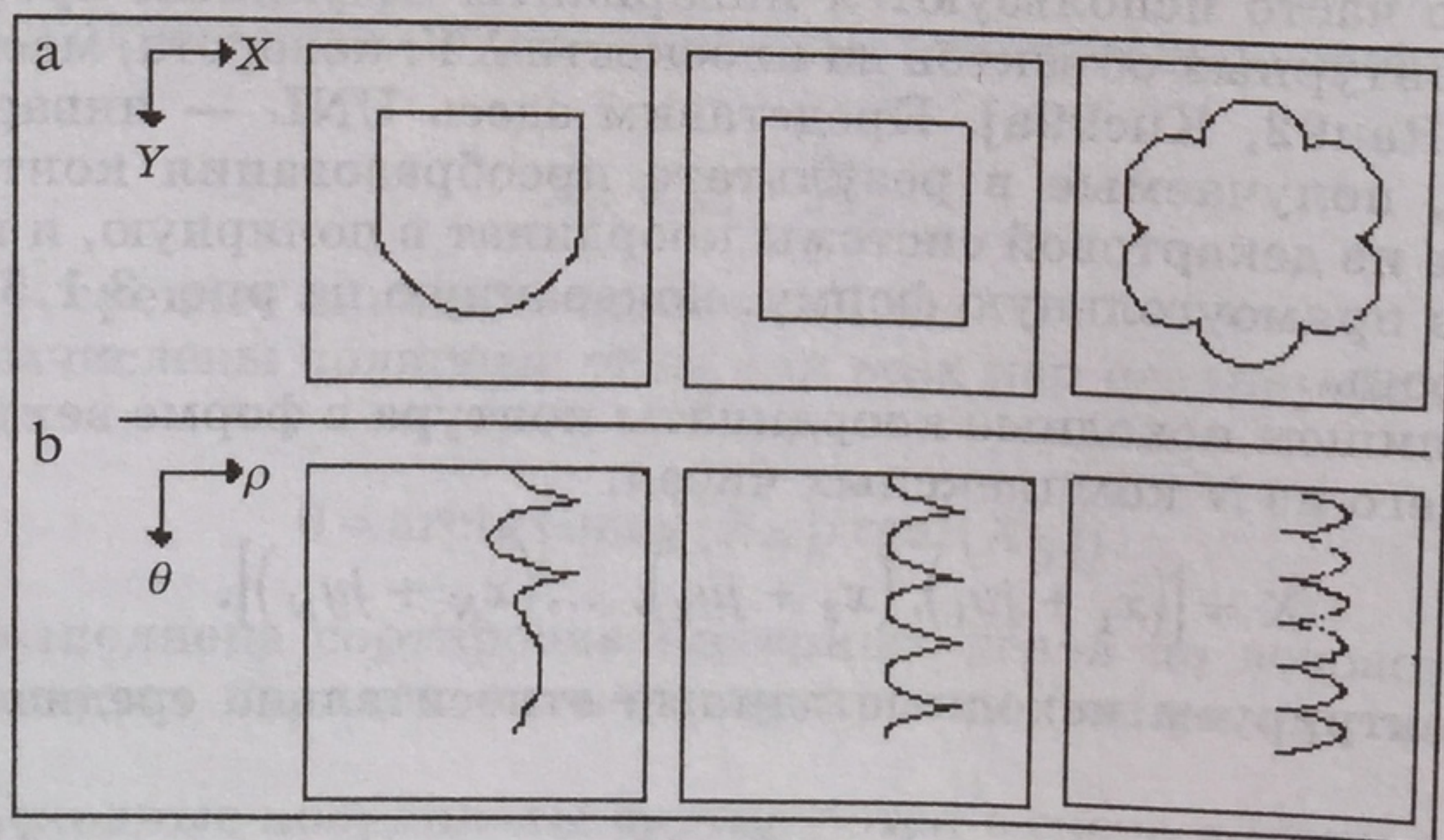


Рис. 3.3.3. Развертка контурного образа с использованием полярных координат: а — исходные образы; б — преобразованные образы



В соответствии с (3.3.11), значения полярного радиуса будут находиться в пределах от 0 до 1. Такая нормировка радиуса позволяет элиминировать масштаб преобразуемых образов. Полярный угол может быть представлен в различных шкалах: либо в диапазоне  $-\pi < \theta < \pi$ , либо в диапазоне от 0 до  $2\pi$ , либо, наконец, в диапазоне от 0 до 1. Если полярные координаты приведены к диапазону от 0 до 1, то пересчет всех значений радиуса и угла  $\theta$  в значения координат соответствующего бинарного UNL-образа (как цифрового изображения) не представляет трудностей. Программа в языке пакета MATLAB, реализующая представленный алгоритм, приведена ниже (Программа 3).

Результат преобразования различных контурных образов по приведенному выше алгоритму представлен на рис. 3.3.3.

### 3.4. ИНВАРИАНТЫ ГЕОМЕТРИЧЕСКИХ ПРЕОБРАЗОВАНИЙ

Один из подходов к решению задач распознавания образов основывается на сравнении каждого нового образа с эталоном и оценке меры подобия между ними. Эталоном в этом случае может быть, например, ранее зарегистрированный образ.

На практике распознаваемый образ и образ-эталон, как правило, отличаются друг от друга масштабом представления, поворотом и сдвигом. Для образов одного и того же класса в большинстве таких случаев исходят из того, что распознаваемый образ появился в результате геометрического преобразования (масштабирования, поворота в плоскости XY и циклического сдвига) образа-эталона. Если последовательно выполнять все возможные геометрические преобразования (повороты, сдвиги и изменения масштаба) эталонного образа и при этом сравнивать результат преобразования с распознаваемым образом то, в конце концов, можно зарегистрировать те параметры преобразований, при которых выступает наивысшее значение меры подобия. Ясно, что такой многоитерационный подход (распознавание через регистрацию), используемый, например, при распознавании отпечатков пальцев, относительно трудоемкий с вычислительной точки зрения.

Другой подход строится не на сравнении самих образов, а на сравнении инвариантов их геометрических преобразований. При этом инвариант эталона вычисляется заранее и хранится в памяти, а сравнение нового образа с эталоном происходит за одну итерацию сразу же после вычисления его инварианта.

На практике распознавания образов применяются инварианты, основанные на моментах, UNL-преобразованиях и преобразованиях Фурье—Меллина, на разностных функциях, оценке функции расстояний и др. [Alf99, Ben98a,b, Che94, Kuc99a, Rau92]. Ниже будут представлены некоторые из перечисленных инвариантов.

**UNL-инварианты.** Первый пример, демонстрирующий неизменность некоторых параметров UNL-инвариантов контурного обра-



|   |                   |                   |
|---|-------------------|-------------------|
| a | 1 1 1 1 1 1 1 1 1 | 4 4 4 4 4 4 4 4 4 |
| b |                   |                   |

Рис. 3.4.1. Исходные образы (a) и их UNL-инварианты (b)

за, подвергнутого геометрическим преобразованиям, приведен на рис. 3.4.1.

Пример взят из практики распознавания рукописных цифр, где широко используется UNL-преобразования [Rau92]. Из рисунка видно, что вычисленные новые образы действительно остаются неизменными независимо от размера цифр и их циклического сдвига в плоскости XY. Поворот же исходного образа в декартовой системе координат на некоторый угол заменяется соответствующим циклическим сдвигом инварианта, представляемого в полярной (развернутой) системе координат. Последнее обстоятельство, как известно, не является препятствием для использования UNL-инвариантов. Для сравнения циклически сдвинутых образов наиболее часто, например, применяется взаимная корреляция и построенная на ней согласованная фильтрация.

Еще один пример UNL-инвариантов представлен на рис. 3.4.2, где в верхней части рисунка приведены пять образов, имеющих различные масштабы, повороты и сдвиги.

Над каждым исходным образом приведены значения угла его поворота в плоскости XY и масштаба его изменения по отношению к первому образу. В нижней части рис. 3.4.2 представлены вычисленные инварианты. Здесь отчетливо заметен циклический сдвиг UNL-преобразованных образов вдоль оси X — развернутого полярного угла. При этом разница в сдвигах между самым левым и самым правым образами составляет половину длины UNL-преобра-

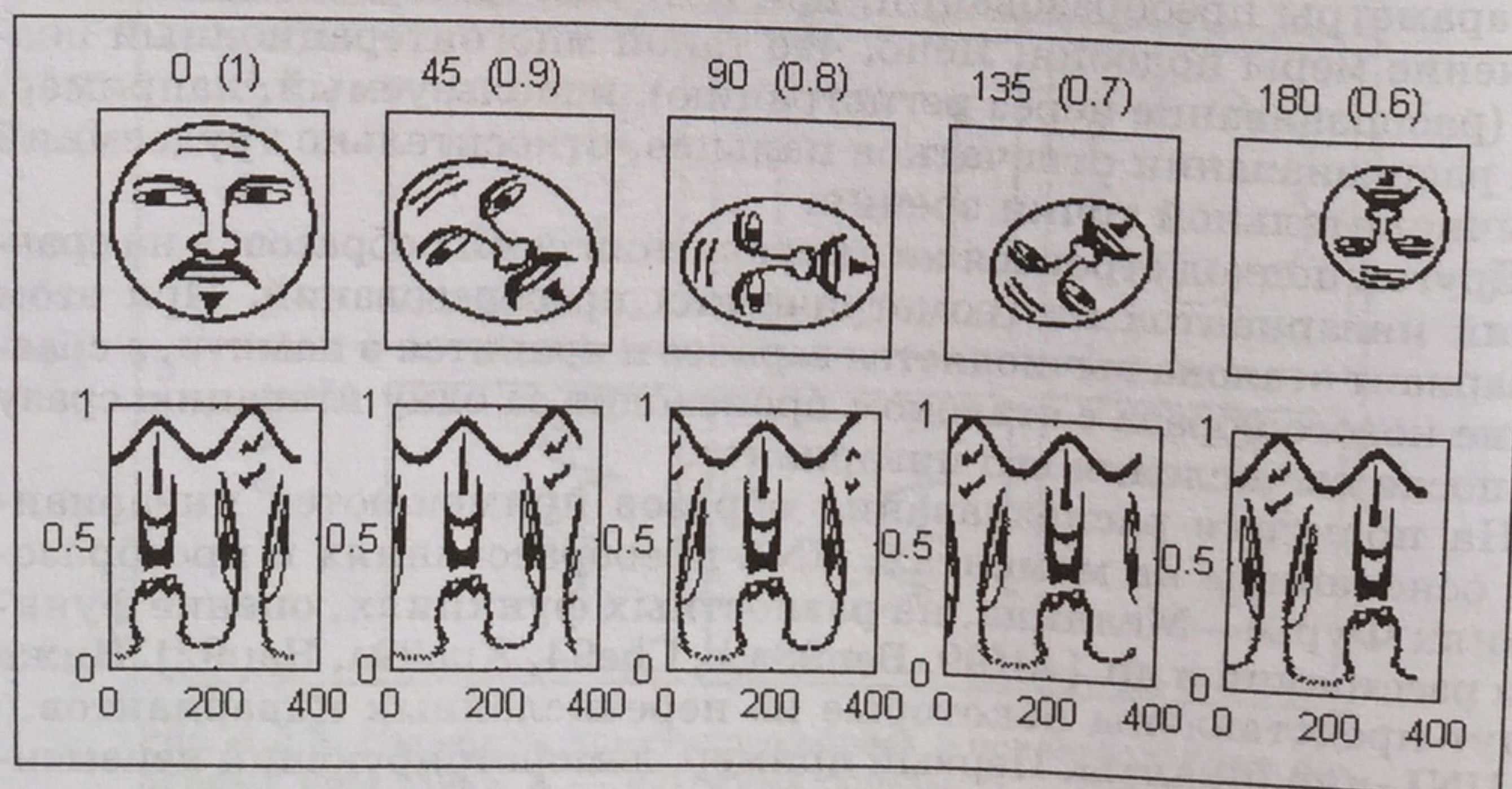


Рис. 3.4.2. Исходные образы и вычисленные для них UNL-инварианты



зованного образа, что строго соответствует разнице в  $180^\circ$  между исходными образами.

Из приведенных на рис. 3.4.1 и 3.4.2 примеров видно, что интерпретация результатов UNL-преобразований и само их вычисление не вызывает трудностей и, следовательно, их можно рекомендовать для применения в задачах распознавания образов. В приложении к биометрике эти преобразования и построенные на них инварианты можно применять в задачах распознавания подписей и других рукописных знаков.

#### Инварианты Affine Arc Length (AAL) и Enclosed Area (EA).

Инварианты AAL и EA для аналитически заданных кривых (контуров) можно определить следующим образом [Alf99, Tie97]:

| Affine Arc Length   | Enclosed Area  | (3.4.1) |
|---|--|---------|
| $\tau = \int_b^a \sqrt{(\dot{x}\ddot{y} - \ddot{x}\dot{y})} dt$ | $\sigma = \frac{1}{2} \int_b^a  x\dot{y} - y\dot{x}  dt$ |         |

Для случая цифровых изображений и контуров, заданных  $M$  координатами, первая и вторая производные заменяются конечными разностями первого и второго порядков:

$$\left. \begin{aligned} \dot{x}(t) &= x(t+1) - x(t); \\ \dot{y}(t) &= y(t+1) - y(t) \end{aligned} \right\} \quad \left. \begin{aligned} \ddot{x}(t) &= \dot{x}(t+1) - \dot{x}(t); \\ \ddot{y}(t) &= \dot{y}(t+1) - \dot{y}(t) \end{aligned} \right\} \quad (3.4.2)$$

для  $t = 1, 2, \dots, M-1$ ;      для  $t = 1, 2, \dots, M-2$ .

С учетом (3.4.2) инвариант EA можно вычислить следующим способом:

$$\sigma(k) = \begin{cases} 0 & \text{для } k = 1; \\ \frac{1}{2} \sum_{t=1}^{k-1} |x(t)y(t-1) - x(t+1)y(t)| & \text{для } k = 2, 3, \dots, M. \end{cases} \quad (3.4.3)$$

Программа в языке пакета MATLAB, предназначенная для вычисления инвариантов AAL и EA, представлена ниже (Программа 4).

Координаты контура (фигуры или образа), для которого определяются инварианты AAL и EA, должны быть заданы в комплексной форме, подобной (3.3.9). Амплитуда вычисленных значений инвариантов AAL и EA нормирована к единице, соответствующие масштабные коэффициенты вычисляются и выводятся как выходные параметры Sa и Se (см. соответствующие строки в конце программы). Число элементов в векторах AAL и EA, представляющих инварианты, должно быть равно числу элементов контура.

Вид инвариантов Affine Arc Length и Enclosed Area показан на рис. 3.4.3. Представленный здесь результат получен по програм-



```
function [ALL, E_A, Sa, Se] = AAL_EA(X)
```

```
% Вычисление инвариантов Affine Arc Length и Enclosed Area.
```

```
% X – вектор координат образа (комплексная форма записи);
```

```
% ALL – функция Affine Arc Length (нормированная к 1);
```

```
% Sa – масштабный коэффициент для Affine Arc Length ;
```

```
% E_A – функция Enclosed Area (нормированная к 1);
```

```
% Se – масштабный коэффициент для Enclosed Area.
```

```
% Выбор координат x и y
```

```
kx = real(X); ky=imag(X);
```

```
% Вычисление первой конечной разности координат
```

```
x_p = [0 diff(kx)]; y_p = [0 diff(ky)];
```

```
% Вычисление второй конечной разности координат
```

```
x_pp = [0 0 diff(diff(kx))]; y_pp = [0 0 diff(diff(ky))];
```

```
% Вычисление инварианта Affine Arc Length
```

```
ALL = cumsum(abs(x_p.*y_pp - x_pp.*y_p).^(1/3));
```

```
Sa = max(ALL); ALL = ALL / Sa;
```

```
% Вычисление инварианта Enclosed Area.
```

```
E_A = cumsum(abs(kx.*y_p - ky.*x_p))/2;
```

```
Se = max(E_A); E_A = E_A/Se;
```

```
% Конец программы
```

ме 4 для образов (первая строка рисунка), аналогичных рассмотренному выше примеру. Всю вторую строку на рис. 3.4.3 занимают инварианты Affine Arc Length, а инварианты Enclosed Area размещены соответственно на третьей строке.

Отмечая независимость полученных результатов от изменений исходных образов (угла поворота, масштаба и сдвига), обратим внимание на некоторую своеобразную «невывразительность» функций, представляющих обсуждаемые инварианты. Последнее обстоятельство, возможно, может затруднить их использование в задачах распознавания при большом числе классов образов.

Кроме того, поскольку вычисление AAL и EA связано с определением конечных разностей (дискретного аналога операции дифференцирования) очень чувствительных к шумам, применение инвариантов AAL и EA требует предварительного сглаживания линий контуров и, если границы их сильно зашумлены, то потребуются их существенная фильтрация [Tie97].

**Инварианты, основанные на матрице расстояний.** Как известно, расстояние (3.1.1) между двумя точками на плоскости не изменяется при их синхронном сдвиге и повороте. Это свойство

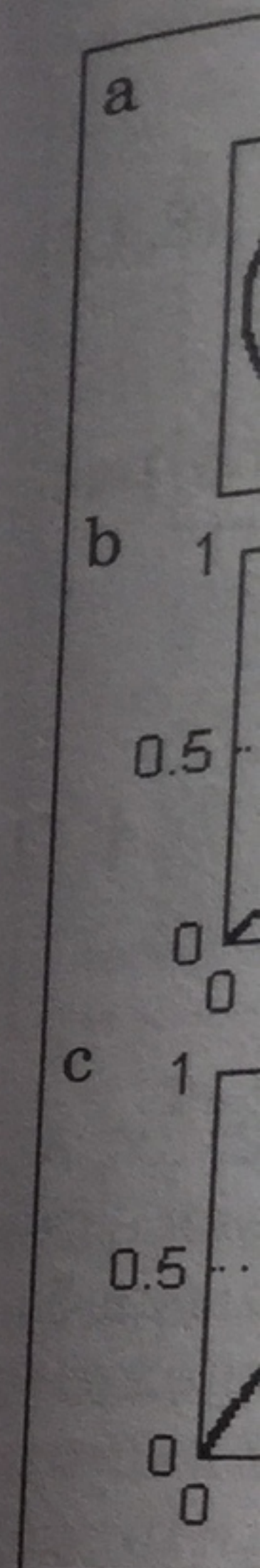


Рис. 3.4.

можно и  
разов на  
Рассм  
расстоян  
образа [К  
тельно к  
представл  
каждой к  
мый конт  
турного о  
ции масш  
наковым.  
Весь пр  
дующим с  
1. Запи  
состояще

2. Цент  
чения:

где  $\bar{X}$  — ср



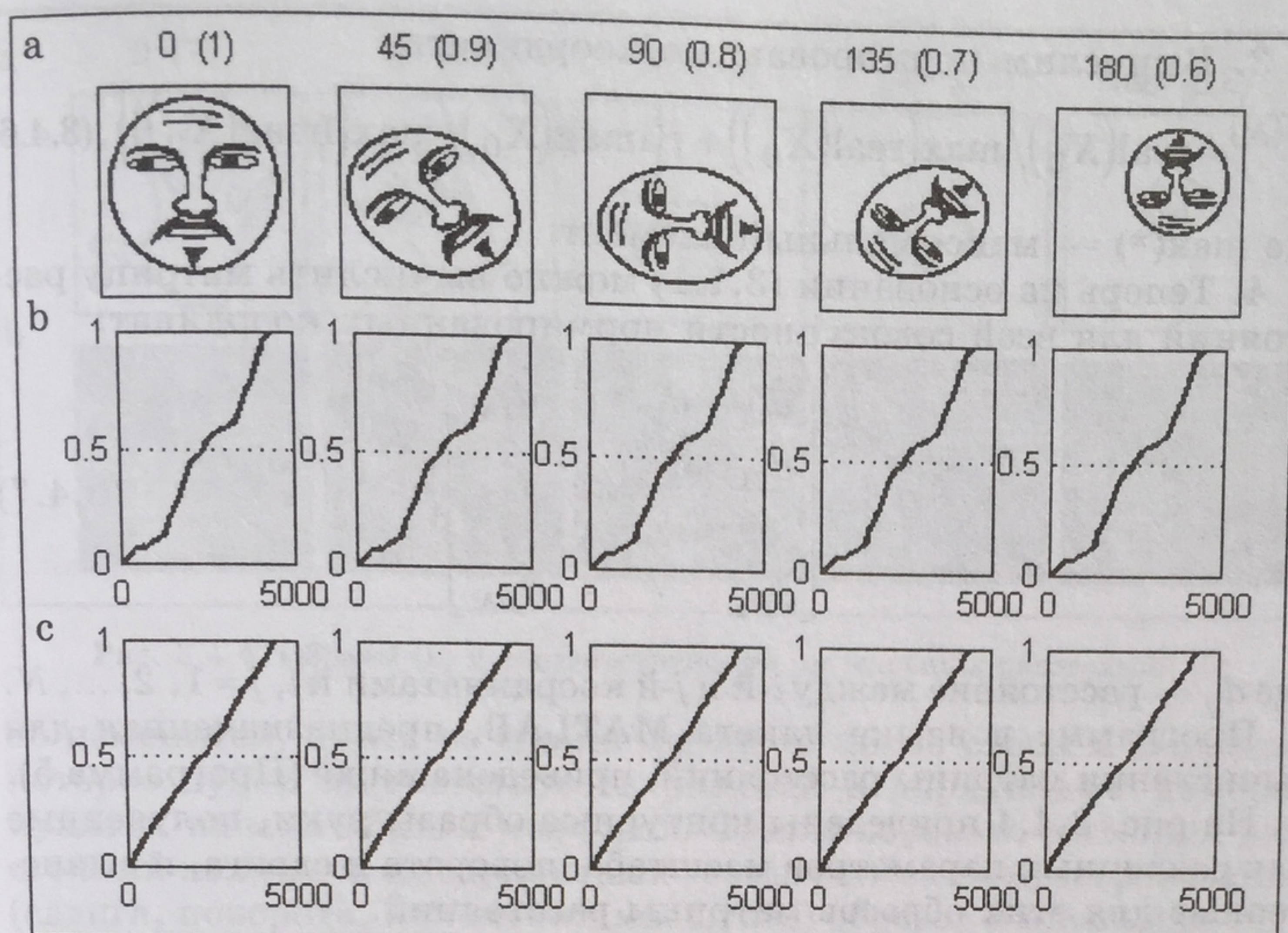


Рис. 3.4.3. Исходные образы (a) и вычисленные для них инварианты Affine Arc Length (b) и Enclosed Area (c)

можно использовать при построении инвариантов контурных образов на бинарных изображениях.

Рассматриваемый здесь класс инвариантов вычисляется как расстояния между нормализованными координатами контурного образа [Кис99а]. Нормализация координат позволяет дополнительно к отмеченному выше свойству элиминировать масштаб представления образа. Нормализация выполняется отдельно по каждой координате. Последнее связано с тем, что рассматриваемый контур может быть получен как проекция на плоскость контурного объекта из трехмерного пространства. При такой проекции масштаб по каждой из координат не всегда сохраняется одинаковым.

Весь процесс вычисления инварианта можно представить следующим образом:

1. Запишем исходные координаты контура в форме вектора  $X$ , состоящего из  $N$  комплексных чисел:

$$X = [(x_1 + jy_1), (x_2 + jy_2), \dots, (x_N + jy_N)]^T. \quad (3.4.4)$$

2. Центрируем исходные данные относительно среднего значения:

$$X_0 = X - \bar{X}, \quad (3.4.5)$$

где  $\bar{X}$  — среднее значение вектора  $X$ .



3. Вычислим нормированные координаты:

$$X^{(N)} = \text{real}(X_0) / \max(\text{real}(X_0)) + j \{ \text{imag}(X_0) / \max(\text{imag}(X_0)) \}, \quad (3.4.6)$$

где  $\max(*)$  — максимальный элемент.

4. Теперь на основании (3.1.1) можно вычислить матрицу расстояний для всей совокупности нормированных координат:

$$D_N = \begin{bmatrix} d_{11} & d_{12} & \cdot & d_{1N} \\ d_{21} & d_{22} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ d_{N1} & \cdot & \cdot & d_{NN} \end{bmatrix}, \quad (3.4.7)$$

где  $d_{ij}$  — расстояние между  $i$ -й и  $j$ -й координатами и  $i, j = 1, 2, \dots, N$ .

Программа в языке пакета MATLAB, предназначенная для вычисления матрицы расстояний, приведена ниже (Программа 5).

На рис. 3.4.4 приведены контурные образы руки, полученные для различных параметров масштаба, поворота и сдвига, и вычисленные для этих образов матрицы расстояний.

Сравнивая между собой полученные матрицы расстояний, можно отметить, что все они одинаковые. Взаимная корреляция между ними равна строго 1. При этом, даже если пункт начала обхода каждого контура будет разным, значение взаимной корреляции

Программа 5

```
function [D, KXY] = inv_D(Kxy);

% Kxy — вектор-столбец координат "X" и "Y" исходного
% контура;
% KXY — вектор-столбец координат "X" и "Y"
% нормализованного контура;
% D — матрица расстояний (инвариант)

% Центрирование и нормирование данных
MN = mean(Kxy); Kxy = Kxy - MN;
Kx = real(Kxy); ky = imag(Kxy);
max_X = max(kx); max_Y = max(ky);
kx = kx/max_X; ky = ky/max_Y;
KXY = kx + j*ky;

% Вычисление матрицы расстояний (инварианта)
D = KXY * ones(1, length(KXY));
D = abs(D - D. ');

% Конец программы
```

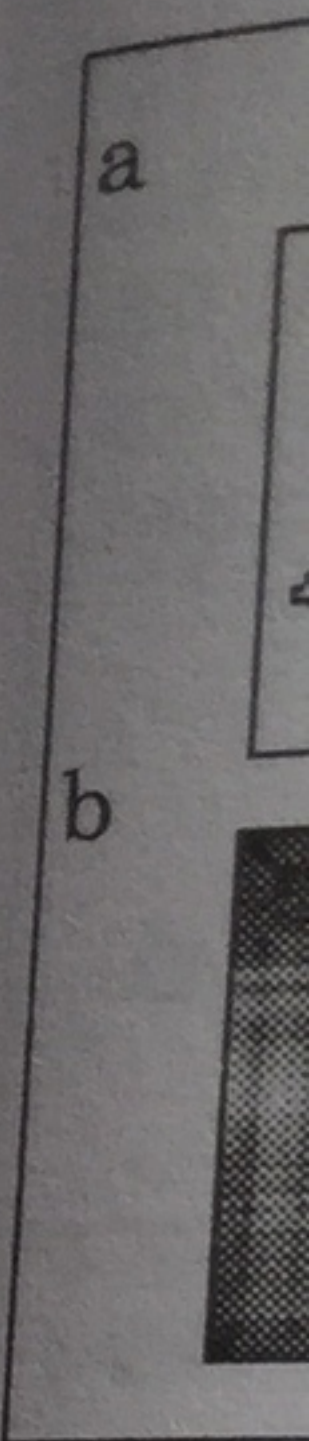


Рис.

не изм  
которо  
пункто  
стояни  
(сдвига

А т

руки с а  
но защ  
нял гра  
нате. Та  
возникн  
образы

В нил  
указаны  
эталона  
корреля  
бии всех

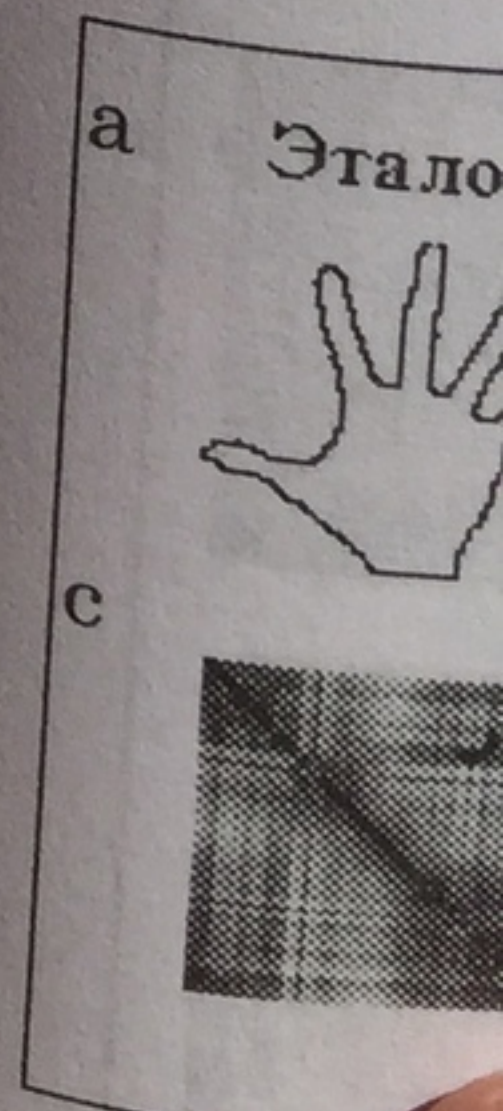


Рис. 3.4.5



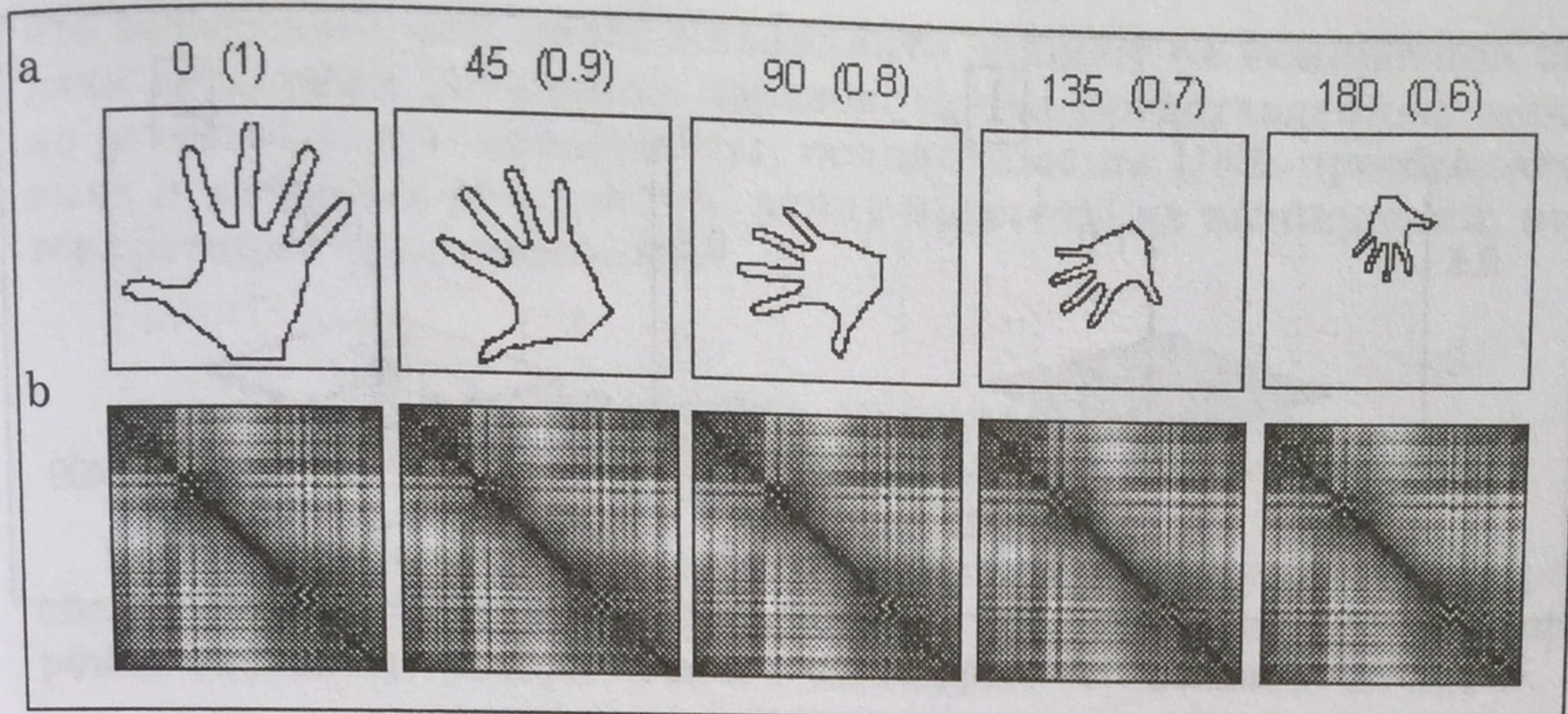


Рис. 3.4.4. Образы (a) и соответствующие им матрицы расстояний (b)

не изменится, а пик ее сдвинется на то число строк и столбцов, которое будет соответствовать разнице в порядковых номерах пунктов начала обхода контуров. Таким образом, матрицы расстояний являются инвариантами геометрических преобразований (сдвига, поворота, изменения масштаба) образов в плоскости  $XU$ .

А теперь представим результаты сравнения образа-эталона руки с аналогичными (рис. 3.4.4) образами, границы которых сильно зашумлены. В данном эксперименте наложенный шум изменял границы контуров в пределах  $\pm 5$  пикселей по каждой координате. Такое или подобное представление контурных образов может возникнуть на практике при обработке реальных данных. Все эти образы представлены на рис. 3.4.5.

В нижней части рис. 3.4.5 под каждой матрицей расстояний указаны оценки взаимной корреляции между матрицей образа-эталона и остальными матрицами. Видно, что значения взаимной корреляции доходят до величины 0,8, что свидетельствует о подобии всех образов между собой.

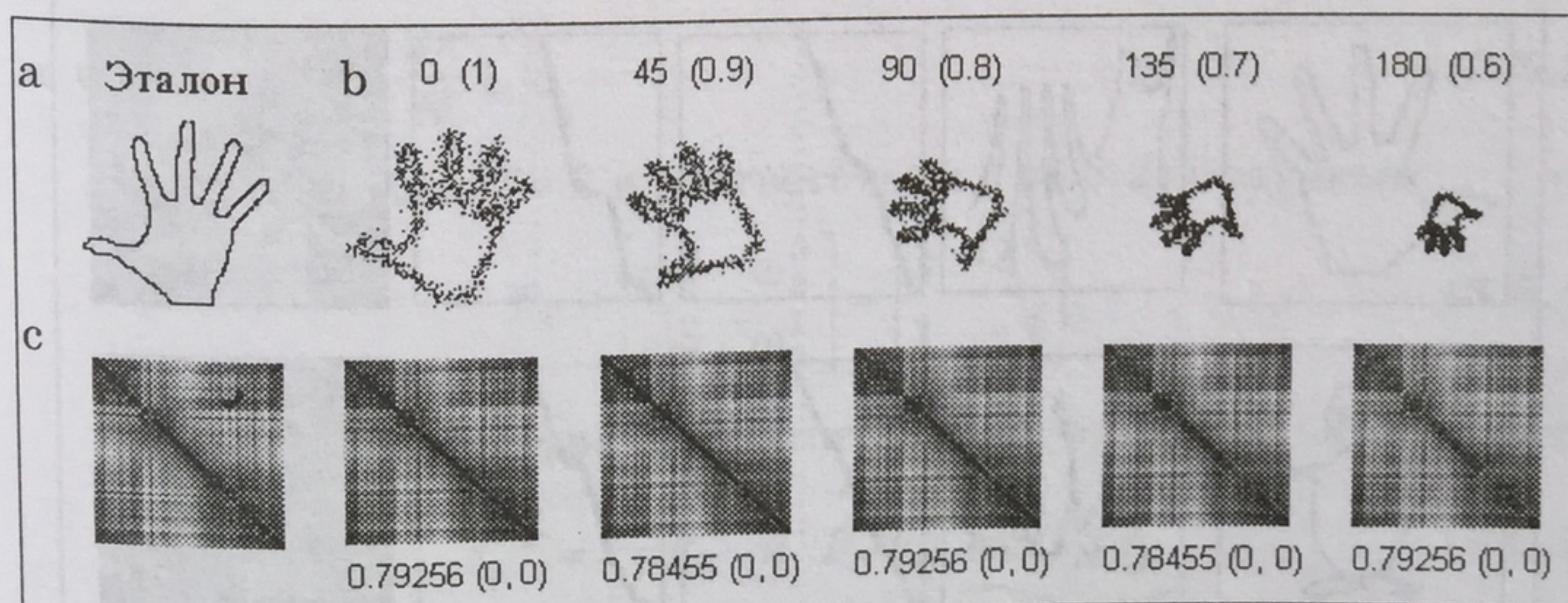


Рис. 3.4.5. Образ-эталон (a), контрольные образы (b) и соответствующие им матрицы расстояний (c), а также корреляция между ними



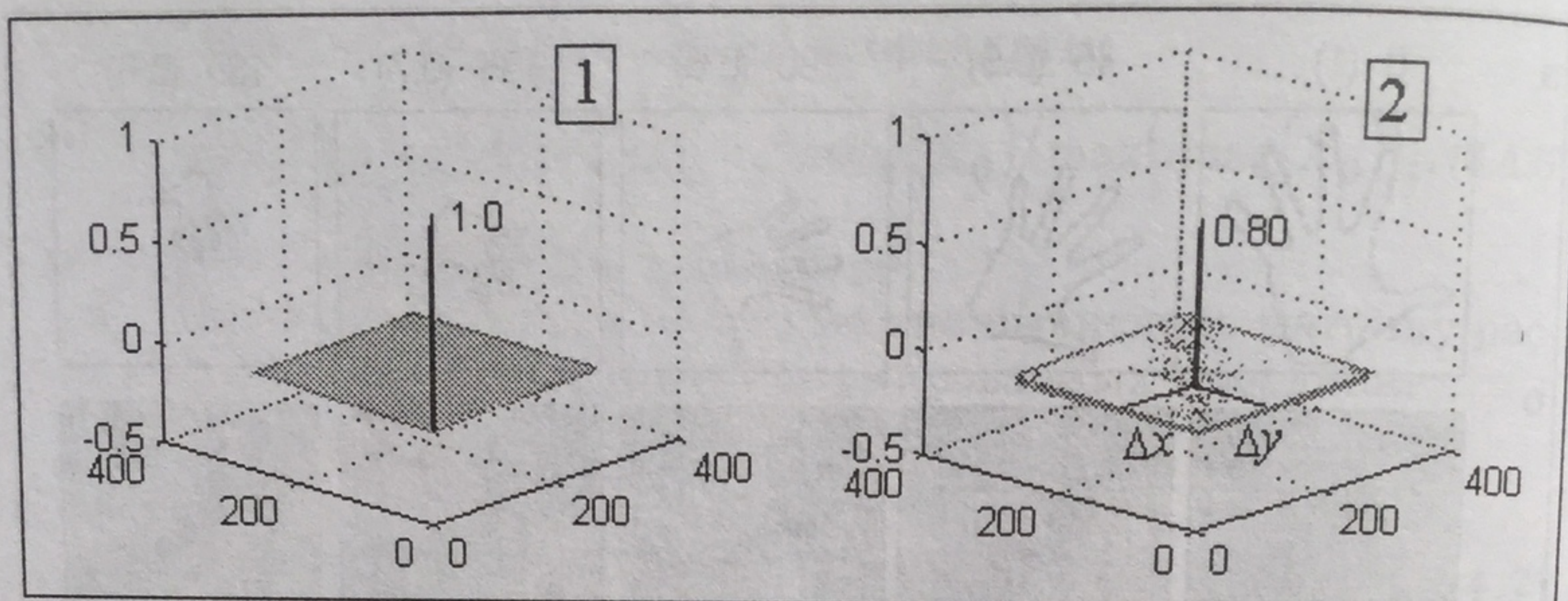


Рис. 3.4.6. Вид двумерной корреляции между образами с рис. 3.4.4 и 3.4.5

Вид взаимной двумерной корреляции для рассмотренного примера показан на рис. 3.4.6. Здесь слева показан график корреляции образов, не подверженных влиянию шума, а справа — двумерная взаимная корреляция образов, подверженных влиянию шума. Сдвиг пика корреляции на величину  $\Delta x = \Delta y$  связан с упомянутым выше различием в выборе начального пункта обхода контуров.

Наконец, отметим, что матрицы расстояний для различных образов будут существенно отличаться, а корреляция между ними будет близка к 0.

Подводя итог сказанному, представим рассмотренные выше инварианты для трех различных образов (рис. 3.4.7). Естественно,

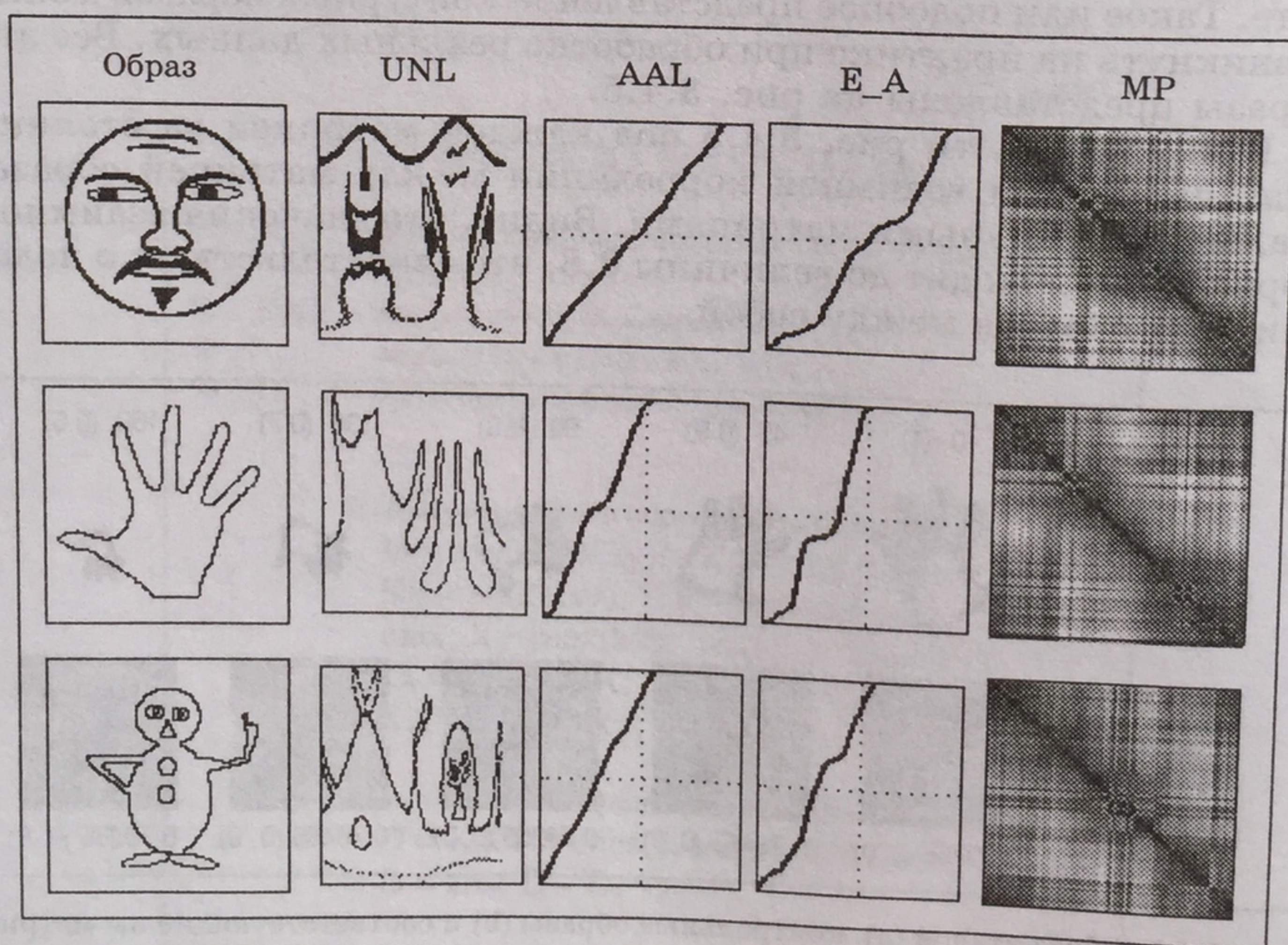


Рис. 3.4.7. Образы и соответствующие им инварианты



что выбор того или иного инварианта зависит от конкретной задачи и условий получения образов. Среди представленных можно рекомендовать инварианты, основанные на UNL-преобразованиях и матрицах расстояний, ввиду простоты их вычислений, интерпретации и применения.

### 3.5. ОБРАБОТКА ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ПРЕОБРАЗОВАНИЙ ФУРЬЕ

В настоящем разделе рассмотрено двумерное дискретное преобразование Фурье (ДПФ) и наиболее часто используемые в цифровой обработке изображений процедуры, основанные на ДПФ.

#### 3.5.1. ИЗОБРАЖЕНИЕ И ЕГО ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Пусть вектор  $X$  представляет некоторый цифровой сигнал во временной области, а вектор  $C$  — этот же сигнал в области преобразования Фурье (спектр сигнала). Тогда прямое и обратное ДПФ (пара ДПФ) могут быть записаны следующим образом [Даг83, Кус98а]:

$$\begin{aligned} C &= F_N X; \\ X &= F_N^{-1} C, \end{aligned} \quad (3.5.1)$$

где

$$F_N = \begin{bmatrix} \omega_N^0 & \omega_N^0 & \omega_N^0 & \dots & \omega_N^0 \\ \omega_N^0 & \omega_N^1 & \omega_N^2 & \dots & \omega_N^{N-1} \\ \omega_N^0 & \omega_N^2 & \omega_N^4 & \dots & \omega_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \omega_N^0 & \omega_N^{(n-1)} & \omega_N^{2(N-1)} & \dots & \omega_N^{(n-1)(n-1)} \end{bmatrix} \quad \begin{array}{l} \text{— матрица преоб-} \\ \text{разования} \\ \text{Фурье} \end{array} \quad (3.5.2)$$

$$\text{и } \omega_M^L = \exp\left(-j \frac{2\pi}{N} L\right). \quad (3.5.3)$$

Например, для  $N = 2$  в соответствии с (3.5.2) получаем

$$F_2 = \begin{bmatrix} \omega_2^0 & \omega_2^0 \\ \omega_2^0 & \omega_2^1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (3.5.4)$$

Матрица  $F_N$  симметричная, унитарная и ортогональная, так что

$$F_N^{-1} = \frac{1}{N} F_N^* \quad \text{и} \quad F_N F_N^* = N I_N. \quad (3.5.5)$$

Пусть теперь нам задана матрица  $X_N$ , которая представляет некоторое цифровое изображение, и матрица  $C_N$ , определяющая



это же изображение в области преобразования Фурье (спектр). Тогда пару двумерных ДПФ (прямое и обратное) можно записать следующим образом:

$$\begin{cases} C_N = F_N X_N F_N; \\ X_N = F_N^{-1} C_N F_N^{-1}. \end{cases} \quad (3.5.6a)$$

Выражение (3.5.6) можно записать также в форме

$$\begin{cases} C(k, l) = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} X(n, m) \exp\left(-j \frac{2\pi}{N} (km + lm)\right); \\ X(n, m) = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} C(k, l) \exp\left(j \frac{2\pi}{N} (km + lm)\right), \end{cases} \quad (3.5.6b)$$

где  $\begin{cases} k = 1, 2, \dots, N; \\ l = 1, 2, \dots, N \end{cases}$  и  $\begin{cases} n = 1, 2, \dots, N; \\ m = 1, 2, \dots, N. \end{cases}$

В выражениях (3.5.6) использован способ представления системы координат, показанный на рис. 3.5.1.

В практике цифровой обработки изображений используется такой способ представления спектра, при котором он становится циклически сдвинутым в центр по осям  $k$  и  $l$ . Результат такого представления спектра показан на рис. 3.5.2. Здесь 1 — исходное изображение; 2, 3 и 4 — абсолютные значения спектра, причем 2 — спектр, представленный в системе координат, показанной на рис. 3.5.1 и определяемый по (3.5.6); 3 — спектр, циклически сдвинутый в центр по оси  $k$  и  $l$ ; 4 — тот же спектр, представленный в форме 3D.

Рассматривая третье изображение на рис. 3.5.2, обратим внимание на его особенность: теоретически здесь должны быть отчетливо видны две оси симметрии спектра. Первая, связанная с тем, что исходное изображение действительное, а спектр определяется по (3.5.6) с использованием симметрической матрицы (3.5.2); вторая ось, определяемая симметрией исходного изображения, например, такого, как образ лица.

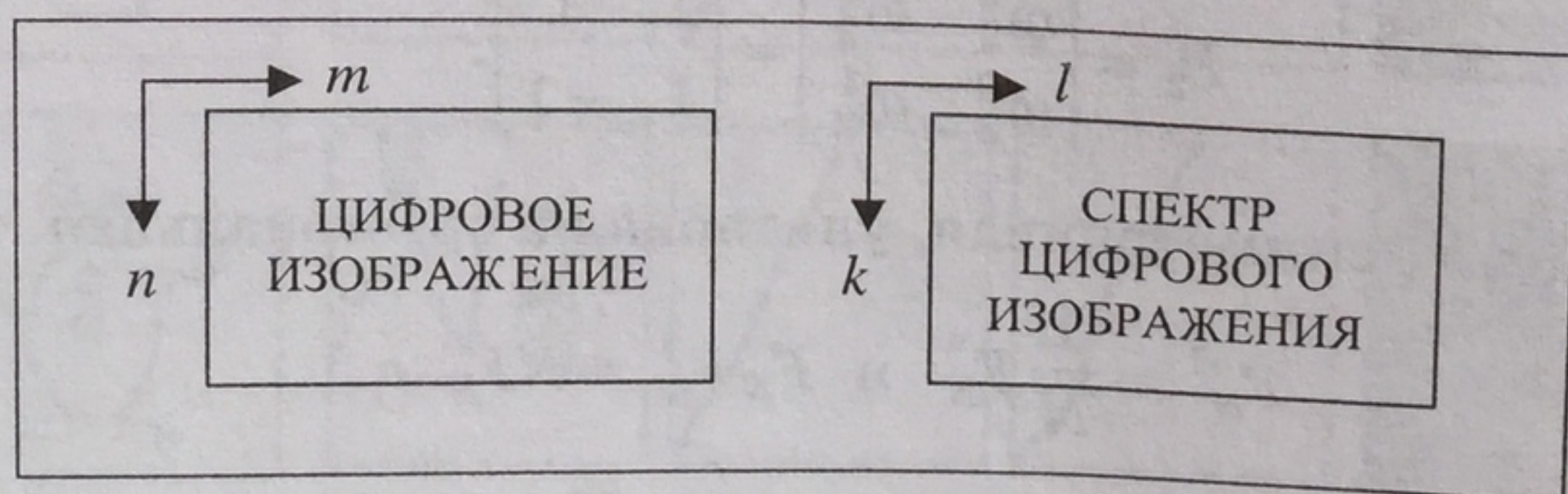


Рис. 3.5.1. Система координат, используемая для представления цифрового изображения и его преобразования Фурье



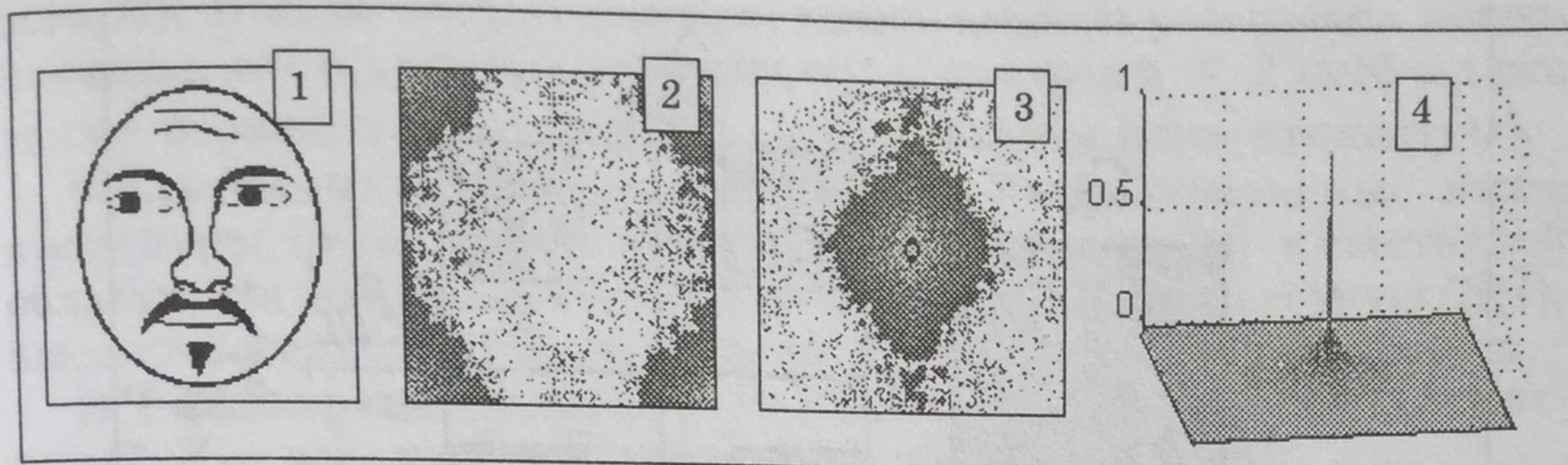


Рис. 3.5.2. Цифровое изображение и формы представления его спектра

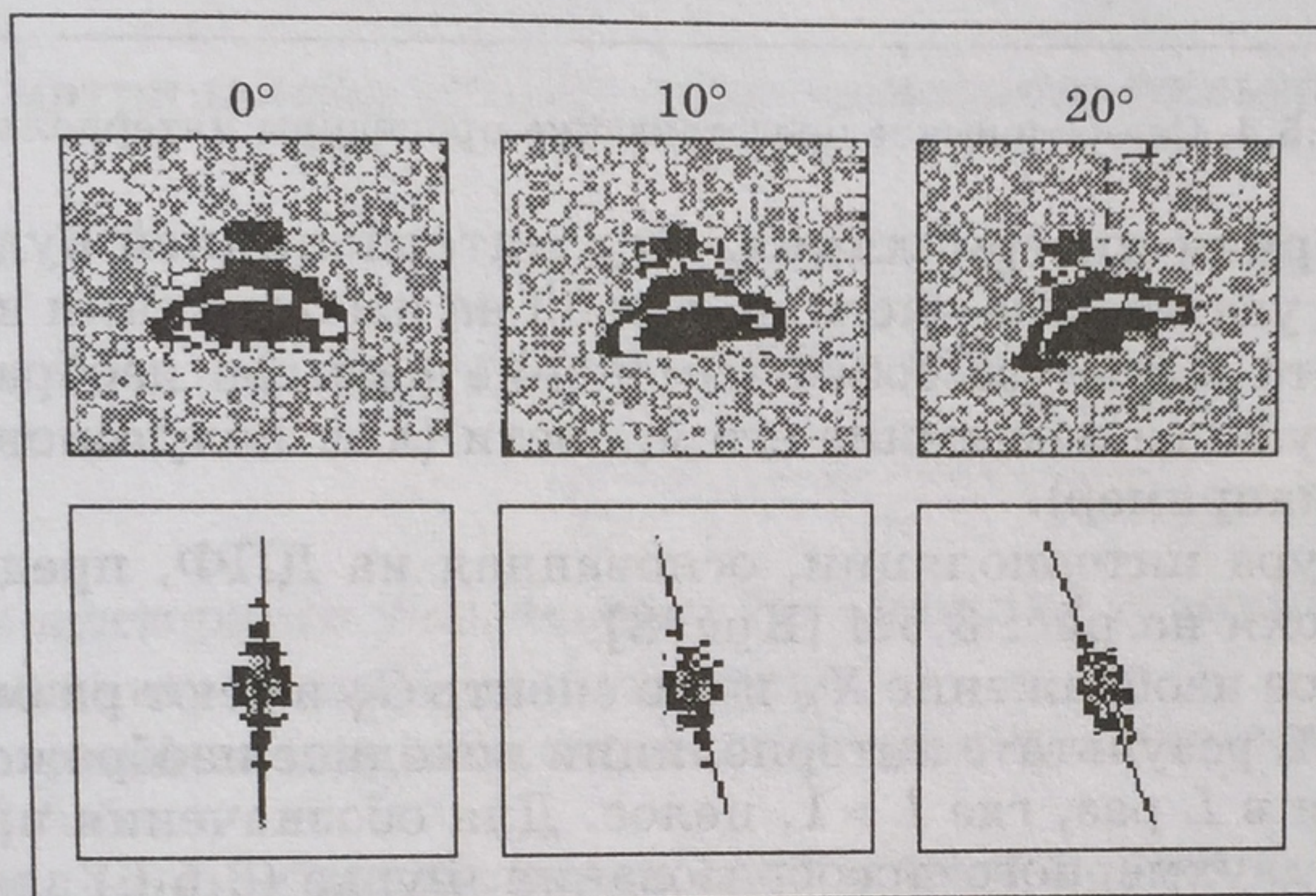


Рис. 3.5.3. Связь между поворотом изображения и его спектром

При этом, если исходное изображение будет повернуто на некоторый угол  $\varphi$  в плоскости  $XU$ , то и его спектр будет также повернут на тот же угол  $\varphi$ . Эту зависимость можно заметить на рис. 3.5.3, где представлены исходные изображения, повернутые в плоскости  $XU$  на разные углы, и соответствующие им спектры, также повернутые на эти же углы.

### 3.5.2. ПРИМЕНЕНИЕ ДИСКРЕТНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Представленные ниже процедуры цифровой обработки изображений основаны на использовании свойств ДПФ (линейности, ортогональности и обратимости ДПФ, существования свойств ДПФ для свернутых и перемноженных сигналов и т. д.). Собственно ДПФ реализуется, как правило, с использованием процедур быстрого преобразования Фурье (БПФ). Как первое, так и второе создают возможности построения простых для понимания и несложных (быстродействующих) для реализации алгоритмов цифровой обработки изображений.



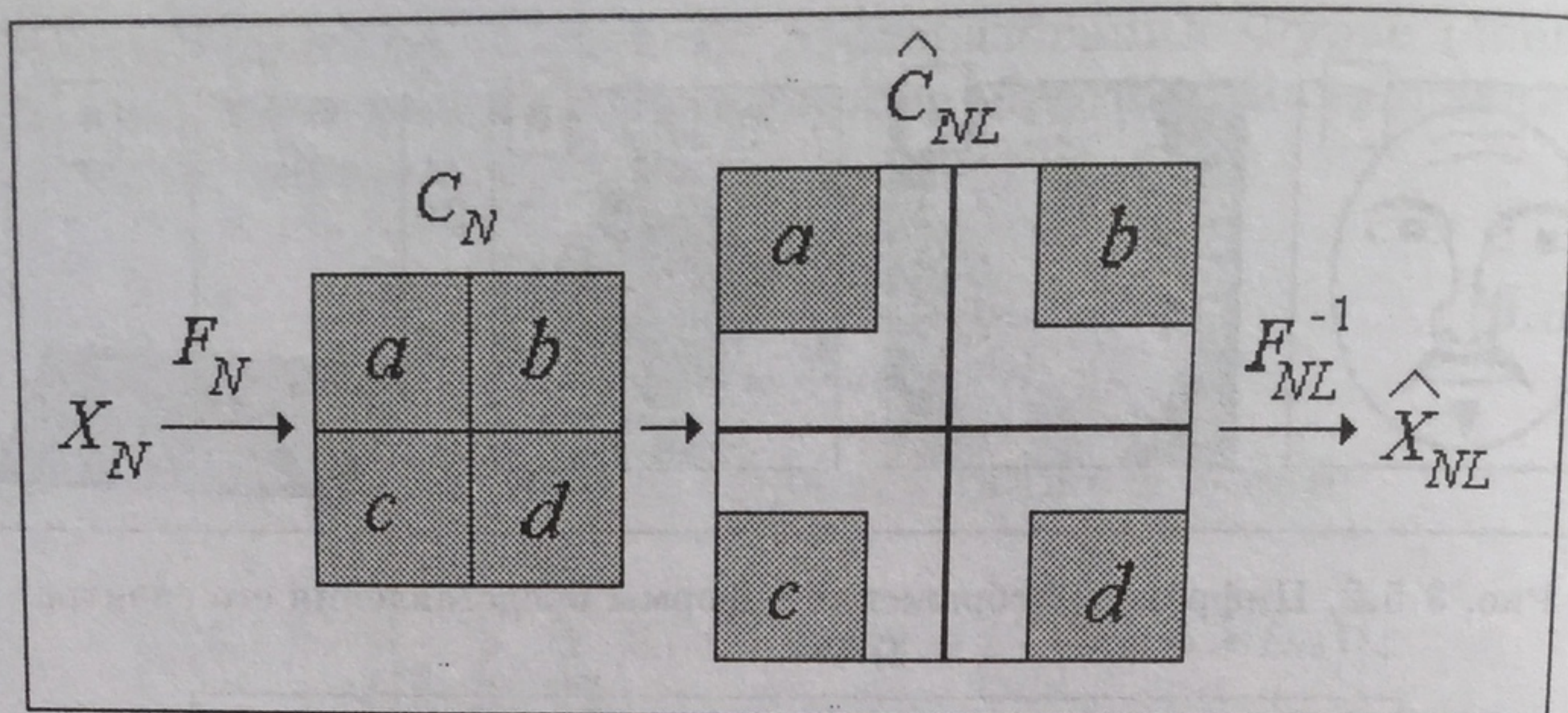


Рис. 3.5.4. Схематическое представление процедуры интерполяции

**Двухмерная интерполяция.** Под интерполяцией будем понимать здесь увеличение числа пикселей на изображении вдоль обеих осей, что влечет за собой изменение размера изображения и соответствующее изменение его яркости (для полутонных изображений, например).

Процедура интерполяции, основанная на ДПФ, представлена схематически на рис. 3.5.4 [Кис98].

Исходное изображение  $X_N$  и его спектр  $C_N$  имеют размер  $N \times N$  пикселей. В результате интерполяции исходное изображение увеличивается в  $L$  раз, где  $L > 1$ , целое. Для обозначения прямого и обратного двухмерного преобразования Фурье (3.5.6) здесь и далее приняты следующие обозначения:

прямое двухмерное ДПФ  $X_N \xrightarrow{F_N} C_N$ ;

обратное двухмерное ДПФ  $\hat{C}_{NL} \xrightarrow{F_{NL}^{-1}} \hat{X}_{NL}$ .

Вся процедура интерполяции в соответствии со схемой 3.5.4 реализуется следующим образом:

- 1) выполняется прямое ДПФ исходных данных;
- 2) матрица спектра  $C_N$  разделяется на четыре подматрицы:  $a$ ,  $b$ ,  $c$ ,  $d$ ;
- 3) формируется нулевая матрица  $\hat{C}_{NL}$  порядка  $NL$ ;
- 4) в углы нулевой матрицы  $\hat{C}_{NL}$  переписываются соответствующие им подматрицы  $a$ ,  $b$ ,  $c$ ,  $d$ ;
- 5) выполняется обратное двухмерное дискретное преобразование Фурье для матрицы  $\hat{C}_{NL}$ , результат которого представляет собой искомую интерполяцию

$$\hat{X}_{NL} = \text{real} \left( F_{NL}^{-1} \hat{C}_{NL} F_{NL}^{-1} \right).$$

При реализации описанной процедуры интерполяции необходимо помнить о спектральных компонентах с номерами  $N/2$  (для чет-

ного  
полн  
такж  
Ф  
прог  
обла  
высо  
Н  
рице

где  $P$   
что

$\times$  — з  
 $\hat{X}_N$  —  
Схе  
на рис  
На  
ции из  
ней гр  
Исх  
ное сл  
кажды

$X_N$

Рис. 3.



ного  $N$ ). В этом случае для получения точного результата интерполяции, спектральные компоненты с номерами  $N/2$  необходимо также перенести в матрицу  $\hat{C}_{NL}$  (на четвертом шаге процедуры).

**Фильтрация в частотной области.** Рассматриваемые ниже процедуры относятся к фильтрации изображений в частотной области. На основе ДПФ можно выполнить низкочастотную (НЧ), высокочастотную (ВЧ), полосовую и режекторную фильтрацию.

**НЧ фильтрация** исходного изображения, представленного матрицей  $X_N$ , реализуется следующим образом [Кис98]:

$$\hat{X}_N = \text{Re} \left\{ F_N^{-1} \left[ (F_N X_N F_N) \odot P_N \right] F_N^{-1} \right\}, \quad (3.5.7)$$

где  $P_N$  — матрица (порядка  $N$ ) низкочастотного фильтра и такая, что

$$P_N = f^T \times f \text{ и } f = \underbrace{\underbrace{11 \dots 1}_{k+1} 000 \dots 000 \underbrace{1 \dots 1}_k}_{N}; \quad (3.5.8)$$

$\times$  — знак векторного умножения;  $k$  — нижняя граница фильтра;  $\hat{X}_N$  — отфильтрованное изображение.

Схема реализации низкочастотной фильтрации представлена на рис. 3.5.5.

На рис. 3.5.6 приведены результаты низкочастотной фильтрации изображения при различных значениях параметра  $k$  — нижней границы НЧ фильтра.

Исходным здесь является выделенное изображение, показанное слева на рис. 3.5.6. Значения параметра  $k$  приведены над каждым отфильтрованным изображением.

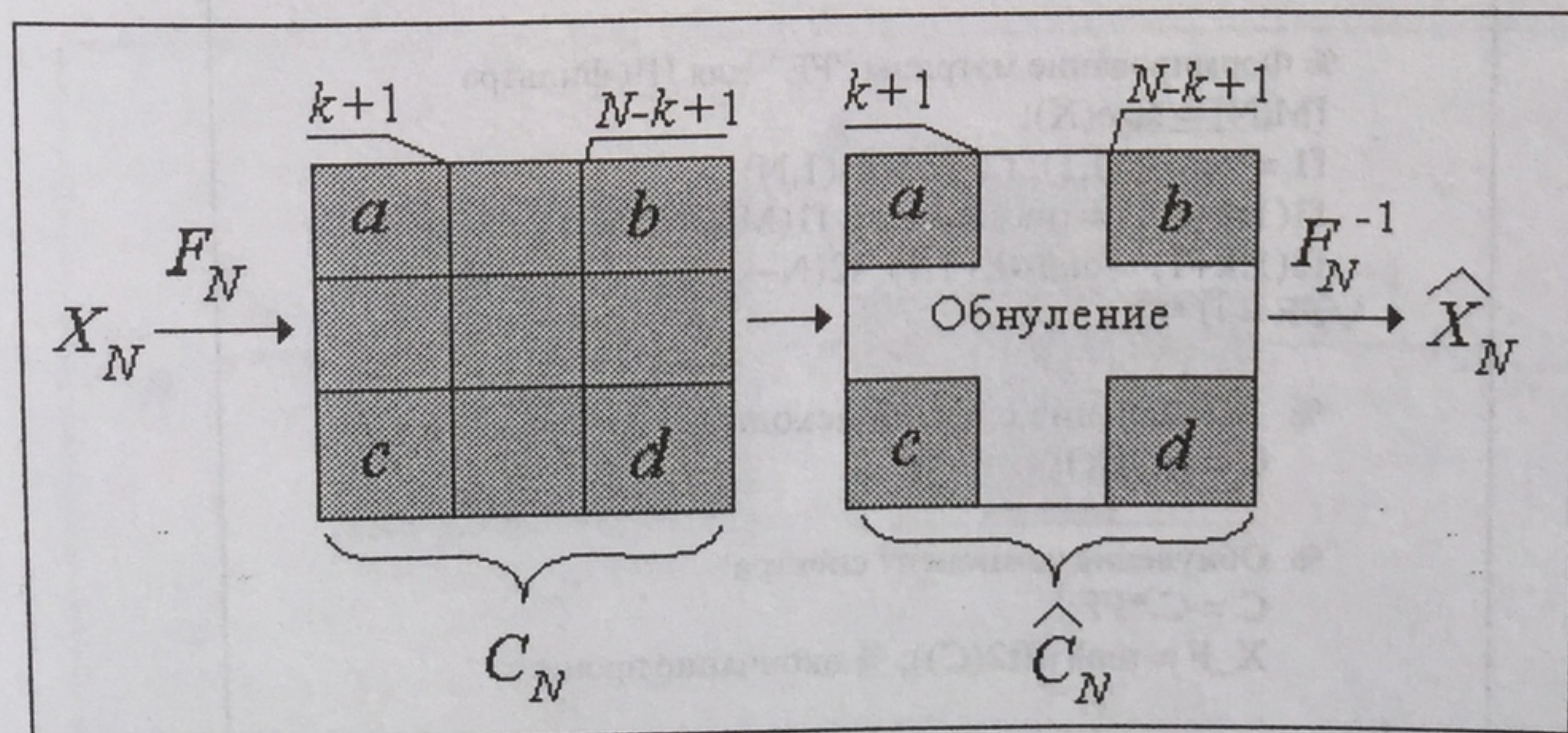


Рис. 3.5.5. Схематическое представление процедуры низкочастотной фильтрации



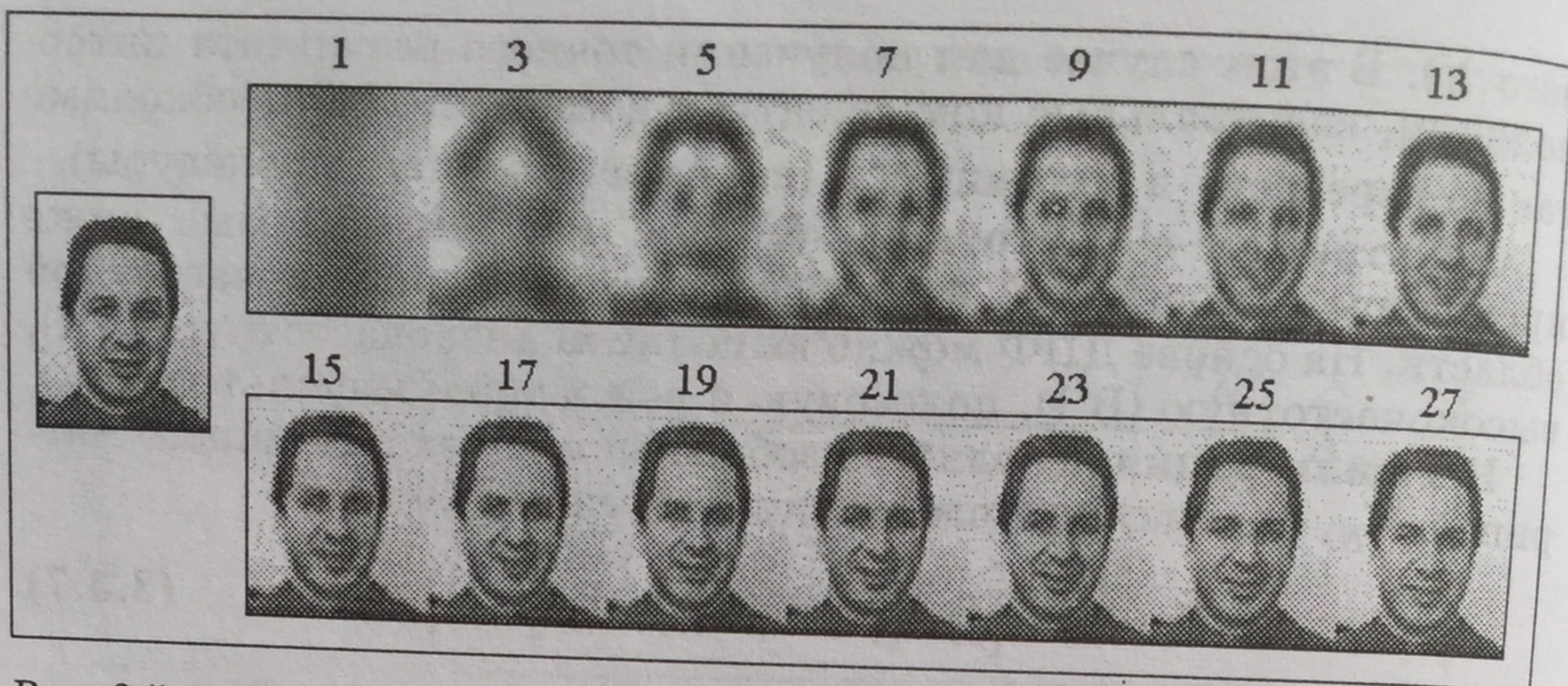


Рис. 3.5.6. Результаты низкочастотной фильтрации в зависимости от границ  $k$  фильтра

Как показывают результаты фильтрации, для удовлетворительного представления образа лица человека достаточно от 20 до 30 пространственных спектральных компонент. Учитывая это обстоятельство, можно значительно сократить размерность исходного пространства признаков, используя, например, в качестве признаков лица его спектральные компоненты.

Программа в языке пакета MATLAB, на основании которой получен представленный на рис. 3.5.6 результат фильтрации, приведена ниже (Программа 6).

Программа 6

```
function X_F = F_2(X,k);

% Низкочастотная фильтрация изображения
% X – исходное изображение
% k – нижняя граница фильтра

% Формирование матрицы "FF" для НЧ фильтра
[M,N] = size(X);
f1 = zeros(M,1); f2 = zeros(1,N);
f1(1:k+1,1) = ones(k+1,1); f1(M-k+1:M,1) = ones(k,1);
f2(1:k+1) = ones(k+1,1); f2(N-k+1:N) = ones(k,1);
FF = f1*f2;

% Вычисление спектра исходного изображения
C = fft2(X);

% Обнуление компонент спектра
C = C.*FF;
X_F = real(ifft2(C)); % окончание процесса

% Конец программы
```



Несомненным достоинством реализации НЧ фильтрации в частотной области (иными словами, через ДПФ) является ее простота и очевидная ясность в выборе параметров фильтра (т. е. ясность в их интерпретации).

**ВЧ фильтрация** цифрового изображения может быть реализована в частотной области двумя представленными ниже алгоритмами [Кис98а].

**А л г о р и т м 1.** Если записать результат НЧ фильтрации некоторого изображения как матрицу  $X_N^{(НЧ)}$ , то ВЧ фильтрация этого же исходного изображения может быть реализована следующим образом [Кис98а]:

$$X_N^{(ВЧ)} = X_N - X_N^{(НЧ)}. \quad (3.5.9)$$

**А л г о р и т м 2.** По этому алгоритму ВЧ фильтрация реализуется аналогично фильтрации НЧ с той лишь разницей, что в матрице  $C_N$  (в спектре, полученном в результате ДПФ исходного изображения) вместо центральной ее части «обнуляются» квадраты в углах этой матрицы.

Вычислительный процесс при этом можно представить как показано на рис. 3.5.7.

**Полосовая и режекторная фильтрация.** Двумерную полосовую и режекторную фильтрацию реализуют обычно с помощью комбинации НЧ и ВЧ фильтрации для различных выбранных параметров  $k$ .

**Вычисление двумерной свертки двух изображений.** В практике цифровой обработки изображений применяют обычно два вида свертки — циклическую и линейную.

Двумерную циклическую свертку  $Z_N$  двух цифровых изображений  $X_N$  и  $Y_N$  можно вычислить следующим образом [Кис98а]:

$$Z_N = \text{Re} \left\{ F_N^{-1} \left[ (F_N X_N F_N) \odot (F_N Y_N F_N) \right] F_N^{-1} \right\}. \quad (3.5.10)$$

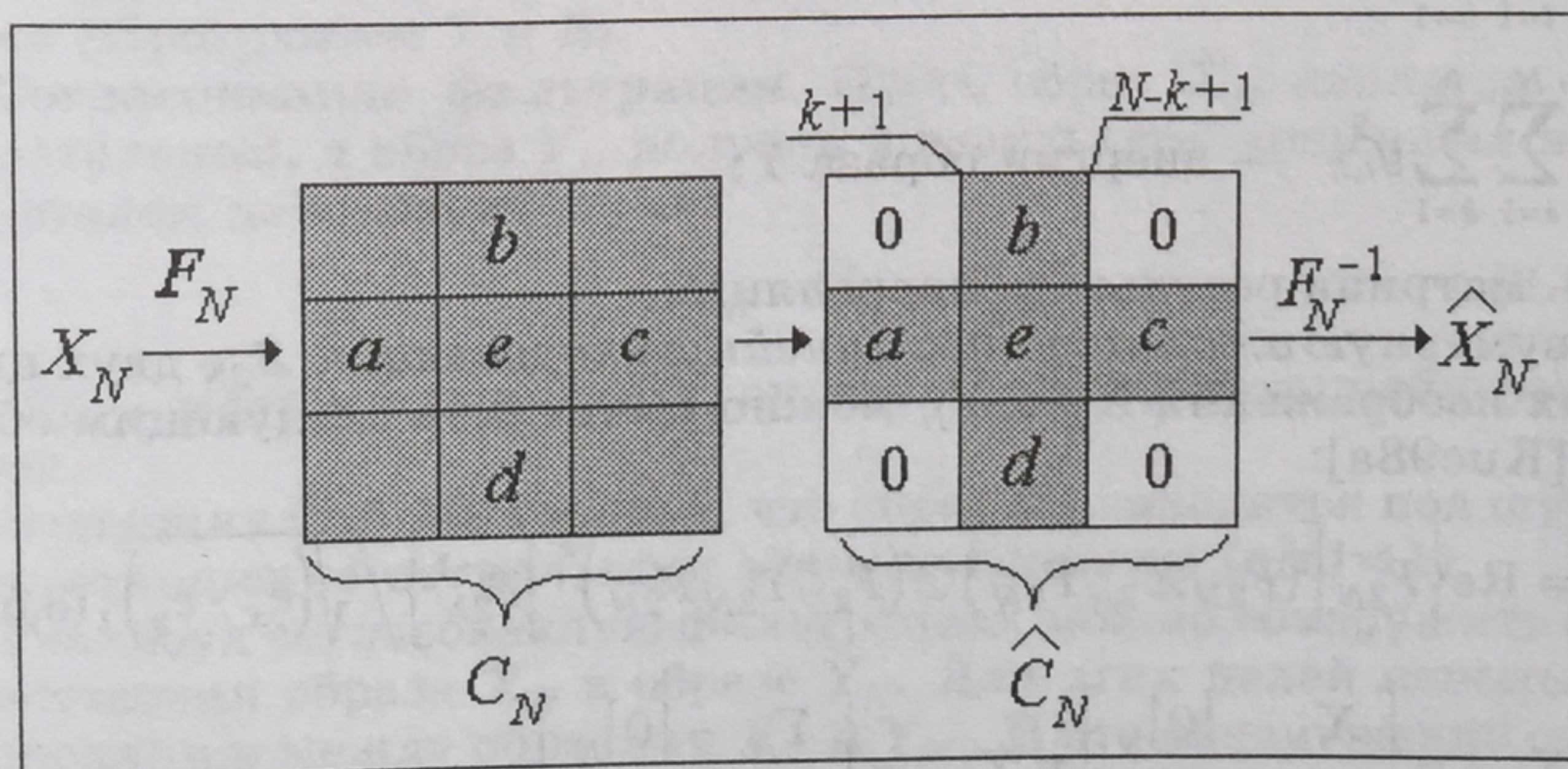


Рис. 3.5.7. Схематическое представление процедуры высокочастотной фильтрации



Здесь и далее использованы обозначения, принятые нами ранее в разд. 3.2 и 3.3. Из (3.5.10) следует, что двумерная циклическая свертка может быть реализована с использованием трех ДПФ и одного поточечного умножения матриц. Из трех ДПФ — два прямых — предназначены для вычисления спектра двух матриц исходных данных, и одно — обратное ДПФ, на его основе формируется результат свертки.

Двумерную линейную свертку  $Z_{2N}$  двух цифровых изображений  $X_N$  и  $Y_N$  можно вычислить следующим образом [Кис98]:

$$Z_{2N} = \text{Re} \left\{ F_{2N}^{-1} \left[ (F_{2N} X_{2N} F_{2N}) \odot (F_{2N} Y_{2N} F_{2N}) \right] F_{2N}^{-1} \right\}, \quad (3.5.11)$$

где  $X_{2N} = \begin{bmatrix} X_N & [0]_N \\ [0]_N & [0]_N \end{bmatrix}$  и  $Y_{2N} = \begin{bmatrix} Y_N & [0]_N \\ [0]_N & [0]_N \end{bmatrix}$ ;

$[0]_N$  — нулевая матрица порядка  $N$ ;  $F_{2N}$  — матрица Фурье порядка  $2N$ ;  $Z_{2N}$  — матрица результата свертки.

**Вычисление двумерной корреляции двух изображений.** Корреляция — один из наиболее распространенных способов проверки соответствия одного образа другому. В практике цифровой обработки изображений наиболее часто применяют следующие виды корреляции: амплитудные циклическую и линейную, а также фазовую.

Двумерную амплитудную циклическую корреляцию  $Z_N$  двух цифровых изображений  $X_N$  и  $Y_N$  можно вычислить следующим образом [Кис98]:

$$R_N = \text{Re} \left\{ F_N^{-1} \left[ (F_N X_N F_N) \odot (F_N Y_N F_N)^* \right] F_N^{-1} \right\} / \sqrt{(e_x \cdot e_y)}, \quad (3.5.12)$$

где  $X_N$  и  $Y_N$  — матрицы исходных данных;

$$e_x = \sum_{i=1}^N \sum_{k=1}^N x_{i,k}^2 \quad \text{— энергия образа } X;$$

$$e_y = \sum_{i=1}^N \sum_{k=1}^N y_{i,k}^2 \quad \text{— энергия образа } Y;$$

$R_N$  — матрица результата корреляции.

Двумерную амплитудную линейную корреляцию  $Z_{2N}$  двух цифровых изображений  $X_N$  и  $Y_N$  можно вычислить следующим образом [Кис98а]:

$$R_{2N} = \text{Re} \left\{ F_{2N}^{-1} \left[ (F_{2N} X_{2N} F_{2N}) \odot (F_{2N} Y_{2N} F_{2N})^* \right] F_{2N}^{-1} \right\} / \sqrt{(e_x \cdot e_y)}, \quad (3.5.13)$$

где  $X_{2N} = \begin{bmatrix} X_N & [0]_N \\ [0]_N & [0]_N \end{bmatrix}$  и  $Y_{2N} = \begin{bmatrix} Y_N & [0]_N \\ [0]_N & [0]_N \end{bmatrix}$ ;



$[0]_N$  — нулевая матрица порядка  $N$ ;  $F_{2N}$  — матрица Фурье порядка  $2N$ ;  $Z_{2N}$  — матрица результата свертки;  $e_x, e_y$  — определены в (3.5.12).

Максимум  $r_{\max}$  двумерной корреляции показывает меру подобия двух изображений (образов) между собой. При этом  $0 \leq r_{\max} \leq 1$ , и, чем ближе  $r_{\max}$  к единице, тем ближе один образ к другому.

Однако адрес максимума корреляции  $A_{\max}$  показывает, на сколько пикселей один образ циклически сдвинут (вверх/вниз или влево/вправо) относительно другого. При этом, если образы не сдвинуты относительно друг друга,  $A_{\max} = (0, 0)$  или  $A_{\max} = (1, 1)$ , что зависит от способа нумерации пикселей на изображениях.

Значение максимума корреляции и его положение — это два основных параметра, используемых при оценке меры соответствия образов друг другу.

Двумерную фазовую корреляцию  $Z_N$  двух цифровых изображений  $X_N$  и  $Y_N$  можно вычислить следующим образом [Кис98а]:

$$R_N^{(\Phi)} = \text{Re} \left\{ F_N^{-1} \left[ \exp \left( j \left( \Phi_N^{(x)} - \Phi_N^{(y)} \right) \right) F_N^{-1} \right] \right\}, \quad (3.5.14)$$

где  $j = \sqrt{-1}$ ;  $\Phi_N^{(x)}$  и  $\Phi_N^{(y)}$  — матрицы фазовых углов для спектров  $C_N^{(x)}$  и  $C_N^{(y)}$ , полученных как результат ДПФ исходных данных  $X_N$  и  $Y_N$  соответственно.

В свою очередь, фазовые матрицы определяются так:

$$\Phi_N^{(x)} = \text{arctg} \left( \frac{\text{Im} \left( C_N^{(x)} \right)}{\text{Re} \left( C_N^{(x)} \right)} \right); \quad \Phi_N^{(y)} = \text{arctg} \left( \frac{\text{Im} \left( C_N^{(y)} \right)}{\text{Re} \left( C_N^{(y)} \right)} \right). \quad (3.5.15)$$

Внешний вид двумерной фазовой корреляции был представлен на рис. 3.4.6 при оценке свойств инвариантов. Программы в языке пакета MATLAB, предназначенные для вычисления взаимной корреляции по процедурам (3.5.12) и (3.5.14) представлены ниже (Программы 7 и 8).

**Согласованная фильтрация.** Пусть образ  $X_N$  является образом-эталоном, а образ  $Y_N$  получен в результате наложения на образ-эталон некоторого шума:

$$Y_N = X_N + S_N, \quad (3.5.16)$$

где  $S_N$  — матрица порядка  $N$ , определяющая некоторый шумовой образ.

В случаях (3.5.16) говорят, что образ  $X_N$  находится под шумом, если отношение сигнал/шум меньше единицы ( $s/\text{ш} < 1$ ).

Реализуя согласованную фильтрацию, можно обнаружить факт присутствия образа  $X_N$  в образе  $Y_N$ . Для этих целей используют корреляцию между образами  $X_N$  и  $Y_N$ . И это обнаружение становится возможным в случаях, когда образы  $X_N$  и  $S_N$  не коррелируют между собой.



Программа 7

```
function [Z, yy, xx, maxZ] = korfun2(X,Y);

% Вычисление амплитудной циклической корреляции
% между двумя образами.
% X и Y – матрицы исходных данных;
% Z – матрица корреляции;
% maxZ – значение максимума корреляции;
% yy, xx – координаты максимума корреляции

% Проверка согласованности размеров данных
if size(X) ~= size(Y) disp(' X ~= Y'); return; end;

% Вычисление спектров
[m,n] = size(X);
Cx = fft2(X)/(m*n); Cy = fft2(Y)/(m*n);

% Вычисление энергии
st1 = std2(Cx); st2 = std2(Cy); mm = st1*st2;

% Вычисление матрицы корреляции
Z = abs(ifft2(Cx.*conj(Cy)))/mm;

% Определение значения и координат максимума
maxZ = max(max(Z));
[yy,xx] = find(Z == maxZ);

% Конец программы
```

Программа 8

```
function [Z, maxZ, yy, xx] = korfaz2(X,Y);

% Вычисление фазовой корреляции между двумя образами.
% X и Y – матрицы исходных данных;
% Z – матрица корреляции;
% maxZ – значение максимума корреляции;
% yy, xx – координаты максимума корреляции

% Проверка согласованности размеров данных
if size(X) ~= size(Y) disp(' X ~= Y'); return; end;

% Вычисление спектров и фазовых углов
[m,n] = size(X);
Cx = fft2(X)/(m*n); Cy = fft2(Y)/(m*n);
Fx = angle(Cx); Fy = angle(Cy);
clear Cx Cy;

% Вычисление матрицы корреляции
jot = sqrt(-1);
Z = real(ifft2(exp(jot*(Fx-Fy)))); clear Fx Fy;

% Определение значения и координат максимума
maxZ = max(max(Z));
[yy,xx] = find(Z == maxZ);

% Конец программы
```



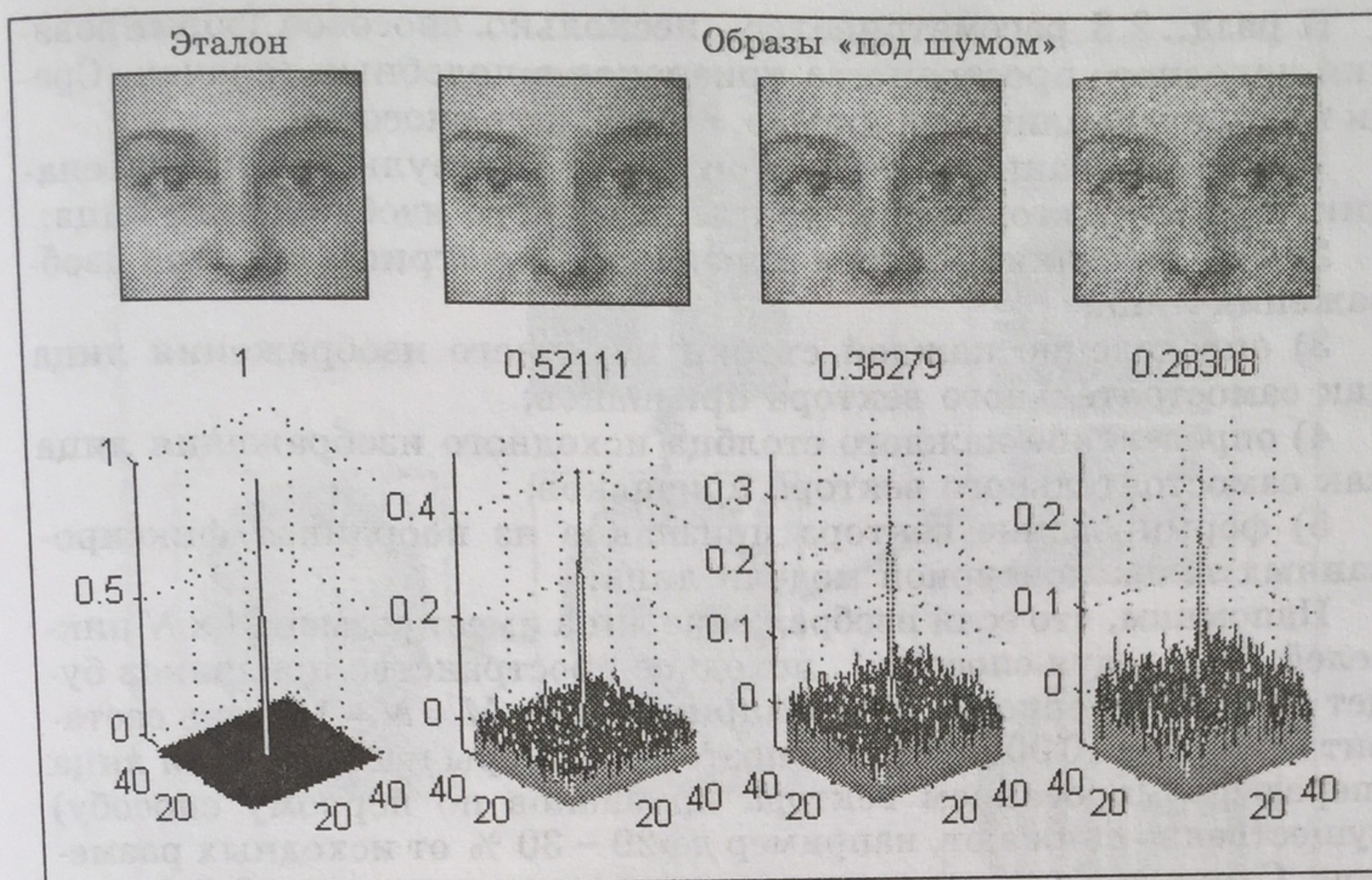


Рис. 3.5.8. Результаты согласованной фильтрации

Пример согласованной фильтрации приведен на рис. 3.5.8, где в верхней части представлены образ-эталон и образы, находящиеся под шумом, а в нижней части показано, как коррелируют эти образы между собой и с образом-эталоном. В данном случае была использована фазовая корреляция. Левая корреляция относится к автокорреляции образа-эталона, поэтому она равна единице.

В задачах распознавания образов оценку факта обнаружения сигнала в шуме определяют по следующему соотношению [Che94]:

$$SSS = \frac{(r_{\max} - M)}{D},$$

где  $M$  — среднее значение корреляции (для всей матрицы корреляции);  $D$  — стандартное отклонение корреляции (для всей матрицы корреляции);  $r_{\max}$  — значение максимума.

При использовании для реализации согласованной фильтрации фазовой корреляции соотношение  $SSS \gg 1$  даже при очень больших амплитудах наложенного шума. Оценивая результат, представленный на рис. 3.5.8, отметим, что поставленная задача согласованной фильтрации успешно решена.

### 3.6. МЕТОД ГЛАВНЫХ КОМПОНЕНТ

В настоящем разделе будет кратко изложена основная идея метода главных компонент (англ. PCA) на примере сокращения исходного пространства признаков в задачах распознавания лиц.



В разд. 2.3 рассматривалось несколько способов формирования исходного пространства признаков в подобных задачах. Среди них упоминались, например, следующие способы:

- 1) формирование вектора признаков в результате конкатенации строк или столбцов матрицы исходного изображения лица;
- 2) формирование вектора признаков из матрицы спектра изображения лица;
- 3) определение каждой строки исходного изображения лица как самостоятельного вектора признаков;
- 4) определение каждого столбца исходного изображения лица как самостоятельного вектора признаков;
- 5) формирование вектора признаков из координат фиксированных точек контурной модели лица.

Напомним, что если изображение лица имеет размер  $M \times N$  пикселей, то, следуя способу 1, исходное пространство признаков будет иметь размерность  $MN$ . Например, для  $M = N = 100$  оно составит величину 10 000. Именно поэтому размеры изображения лица (перед формированием вектора признаков по первому способу) существенно снижают, например до 20 – 30 % от исходных размеров. Схематически этот процесс представлен на рис. 3.6.1, где цифрой 0 обозначено исходное изображение лица, а цифрой 1 — то же изображение, но уменьшенное в размерах и предназначенное для конкатенации его строк (или столбцов) в вектор. Цифрой 2 обозначен спектр исходного изображения, из которого формируется вектор признаков с учетом симметрии спектра и важности только первых 20 пространственных компонент. Изображения, отмеченные цифрами 3 и 4, соответствуют третьему и четвертому способам получения векторов исходного пространства признаков. И, наконец, изображение, отмеченное цифрой 5, определяет последний способ формирования вектора признаков и соответствующего ему пространства признаков.

Размерность пространства признаков, равная числу компонент в векторе признаков, которая остается в результате выполнения указанных выше действий для каждого из способов, приведена в табл. 3.6.1.

Признаки, получаемые любым из представленных выше способов, имеют две особенности: размерность пространства признаков остается все еще значительной (см. табл. 3.6.1), а ранг отдельных признаков (или их важность для представления всех и каждого отдельного исходного изображения) еще не определен. Обе эти особенности, естественно, усложняют, а порой и затрудняют решение проблемы распознавания лиц.

Анализ главных компонент позволяет решить две главные задачи:

- 1) определить ранг признаков;
- 2) существенно сократить размерность пространства признаков.

Покажем теперь, в чем заключается основная идея метода главных компонент, на примере сокращения пространства признаков, определяемого по способу 1 (см. рис. 3.6.1).



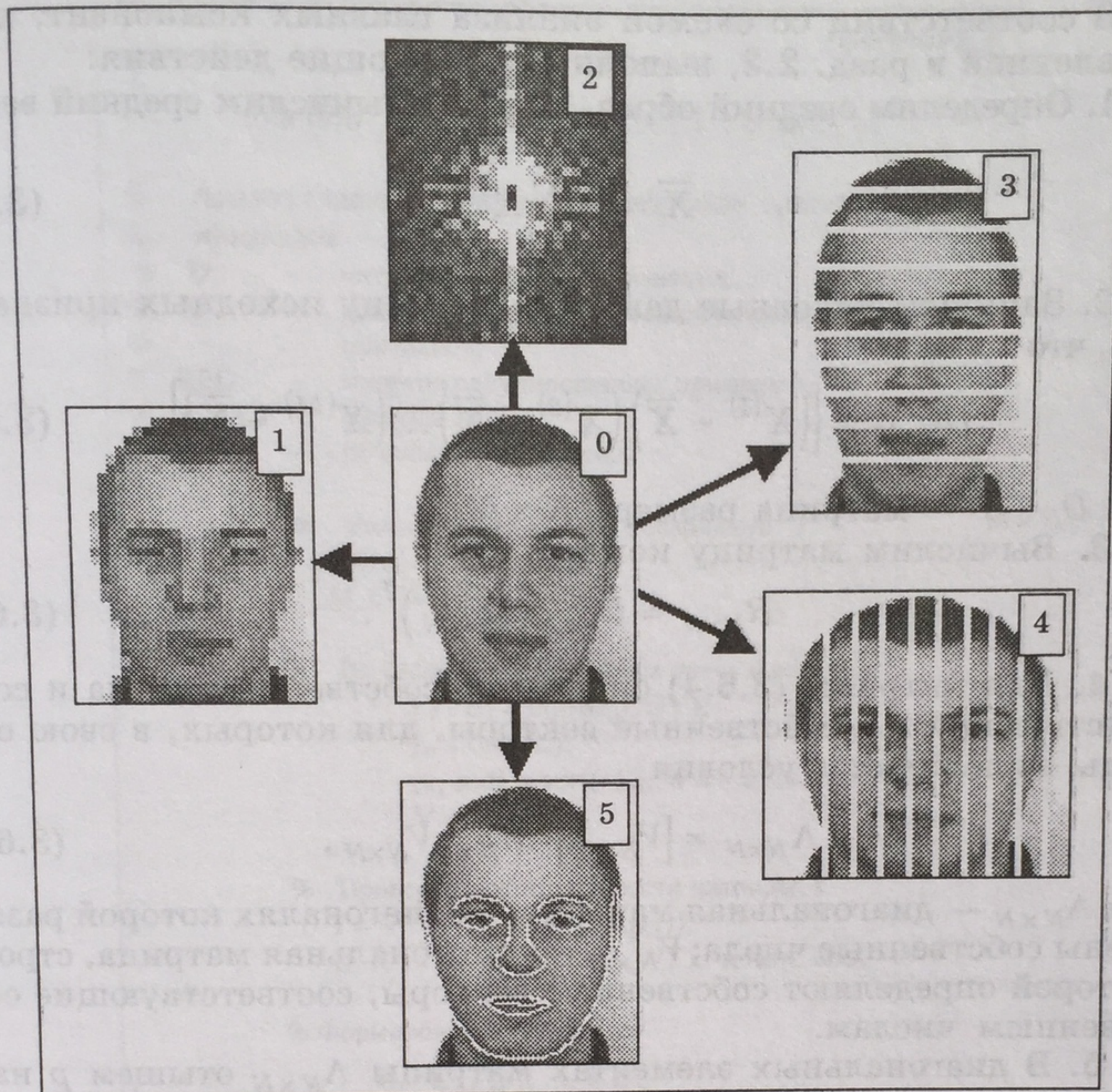


Рис. 3.6.1. Изменения исходного изображения в процессе формирования исходного пространства признаков

Следуя общепринятой терминологии задач распознавания, представим метод главных компонент, используя понятие образа как отдельного элемента класса образов и понятие вектора признаков, определяющего каждый образ в классе.

Итак, исходным для нас является представление каждого образа вектором признаков

$$X^{(i)} = [x(1, i), x(2, i), \dots, x(N, i)]^T, \quad (3.6.1)$$

где  $i$  — номер образа ( $i = 1, 2, \dots, M$ );  $x(n, i)$  — значение признака;  $n$  — номер признака ( $n = 1, 2, \dots, N$ );  $N$  — размерность пространства признаков (число признаков).

Таблица 3.6.1

Размерность пространства признаков

| Способ выбора признаков  | 1        | 2          | 3       | 4       | 5          |
|--------------------------|----------|------------|---------|---------|------------|
| Размерность пространства | 600–1000 | $\geq 400$ | 100–200 | 100–200 | $\geq 150$ |



В соответствии со схемой анализа главных компонент, представленной в разд. 2.3, выполним следующие действия:

1. Определим средний образ, для чего вычислим средний вектор

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X^{(i)}. \quad (3.6.2)$$

2. Запишем исходные данные в матрицу исходных признаков так, что

$$D_{N \times M} = \left[ \left( X^{(1)} - \bar{X} \right) \left( X^{(2)} - \bar{X} \right) \dots \left( X^{(M)} - \bar{X} \right) \right], \quad (3.6.3)$$

где  $D_{N \times M}$  — матрица размера  $N \times M$ .

3. Вычислим матрицу ковариации

$$R_{N \times N} = D_{N \times N} (D_{N \times N})^T. \quad (3.6.4)$$

4. Для матрицы (3.6.4) определим собственные числа и соответствующие им собственные векторы, для которых, в свою очередь, выполняются условия

$$\Lambda_{N \times N} = [V_{N \times N}]^T R_{N \times N} V_{N \times N}, \quad (3.6.5)$$

где  $\Lambda_{N \times N}$  — диагональная матрица, на диагоналях которой размещены собственные числа;  $V_{N \times N}$  — ортогональная матрица, строки которой определяют собственные векторы, соответствующие собственным числам.

5. В диагональных элементах матрицы  $\Lambda_{N \times N}$  отыщем  $p$  наибольших собственных чисел, причем  $p \ll N$  (именно здесь как бы определяется ранг признаков).

6. Из матрицы  $[V_{N \times N}]^T$  выберем  $p$  строк, соответствующих  $p$  наибольшим собственным числам, из которых сформируем матрицу  $F_{p \times N}$  редукции признаков (матрицу преобразования Карунена—Лоэва).

7. Для каждого образа из матрицы исходных признаков выполним преобразования Карунена—Лоэва, которое может быть представлено в форме

$$Y_{p \times M} = F_{p \times N} D_{N \times M}, \quad (3.6.6)$$

где  $Y_{p \times M} = [Y^{(1)} Y^{(2)} \dots Y^{(M)}]$  — матрица признаков в редуцированном пространстве;  $Y^{(i)}$  — образ в редуцированном пространстве признаков.

Отметим, что в соответствии с (3.6.6) каждый образ представлен теперь только  $p$  новыми признаками в редуцированном пространстве. Если использовать равенство Парсеваля для оценки энергии такого образа, то окажется, что новые компоненты описывают значительно большую часть энергии, чем оставшиеся (неучтенные) компоненты. Эти  $p$  новых компонент и называют главными компонентами.



```

function [RED, A_KLT, L] = pca(D, p)

% Анализ главных компонент и редукция пространства
% признаков
% D — матрица исходных признаков;
% p — размерность редуцированного пространства
% признаков;
% RED — матрица редуцированных признаков;
% A_KLT — матрица KLT;
% L — «p» собственных числа;

% Вычисление матрицы ковариации
[N,M] = size(D);
M_COV = (D*D')/M;

% Вычисление собственных чисел и векторов
[V, L] = eig(M_COV);
L = diag(L);
[a, ind] = sort(-L); a = -a; L = a;
V = V(:,ind);

% Проверка ортогональности матрицы V
I = V*V'; R = sum(diag(I));
if R ~ N disp('???_1_???'); return; end;

% Формирование результата
L = L(1:p); L = L';
A_KLT = V(:,1:p)';
RED = A_KLT*D;

% Конец программы

```

Размерность редуцированного пространства признаков равна теперь также  $p$ , что меньше исходной размерности в  $N/p$  раз. Число всех признаков равно  $pM$ , что меньше первоначального числа также в  $N/p$  раз.

Программа в языке пакета MATLAB, предназначенная для анализа главных компонент, вычисления матрицы преобразования Карунена—Лоэва и редукции пространства признаков, приведена выше (Программа 9).

### 3.7. ПРЕОБРАЗОВАНИЕ ФУРЬЕ—МЕЛЛИНА

Выше в разд. 3.1, 3.3 и 3.4 мы отмечали возможность представления исходных изображений в развернутой полярной системе координат, а также анализировали уникальные инвариантные свойства получаемых при этом новых изображений. Эти же подходы



могут быть применены не к исходным изображениям, а к их спектрам, что также создает основу для получения аналогичных инвариантов плоских аффинных преобразований изображений. В приложении к спектрам эти преобразования определены в [Che94] как преобразования Фурье—Меллина.

Однако необходимо отметить, что в обоих рассматриваемых случаях цифровые изображения являются в некоторой степени специфическими — это изображения, в поле рамки которых всегда выделен некоторый объект (один или несколько). Поэтому аффинные (геометрические) преобразования — повороты, сдвиги и масштабирование — касаются именно этих выделенных на всем изображении объектов. Примеры таких изображений — см. на рис. 3.4.1 – 3.4.5.

В приложении к биометрическим технологиям распознавания в [Ben98a,b] на основе преобразований Фурье—Меллина, определяются инварианты поворотов (влево/вправо) и наклонов (вверх/вниз) лица человека.

Идея вычисления Фурье—Меллина инвариантов (ФМИ) состоит в следующем.

1. Для исходных данных, представленных матрицей  $X_N$ , вычисляется модуль спектра двумерного ДПФ:

$$C_N = |F_N X_N F_N|. \quad (3.7.1)$$

2. Матрица спектра (3.7.1) циклически сдвигается по строкам и столбцам так, что

$$C_N^{(\Leftrightarrow)} = \left[ I_N^{(N/2 \rightarrow)} \right]^T C_N I_N^{(N/2 \rightarrow)}, \quad (3.7.2)$$

где  $I_N^{(N/2 \rightarrow)}$  — единичная матрица порядка  $N$  с циклически сдвинутыми на  $N/2$  позиций вправо столбцами [Даг83].

Пример получаемого по (3.7.2) спектра показан на рис. 3.5.2. Отметим, что модуль спектра (3.7.1) является инвариантом циклического сдвига некоторого образа в поле рамки всего изображения, а представление спектра в форме (3.7.2) позволяет легко преобразовать спектр  $|C_N^{(\Leftrightarrow)}|$  из декартовой системы координат в развернутую полярную систему  $|C_N^{(p,\theta)}|$ . Как отмечалось выше, при таком представлении спектра любой поворот исходного образа в плоскости  $XU$  (и в поле рамки исходного изображения) заменяется циклическим сдвигом спектра вдоль оси  $\theta$  развернутого полярного угла.

Если при представлении развернутой полярной системы полярный радиус определять в логарифмической шкале, то полученный спектр будет представлять собой спектр Фурье—Меллина. Именно этот спектр и есть инвариант масштабных изменений (практически до 2 – 3 раз) исходного образа в поле рамки всего изображения.



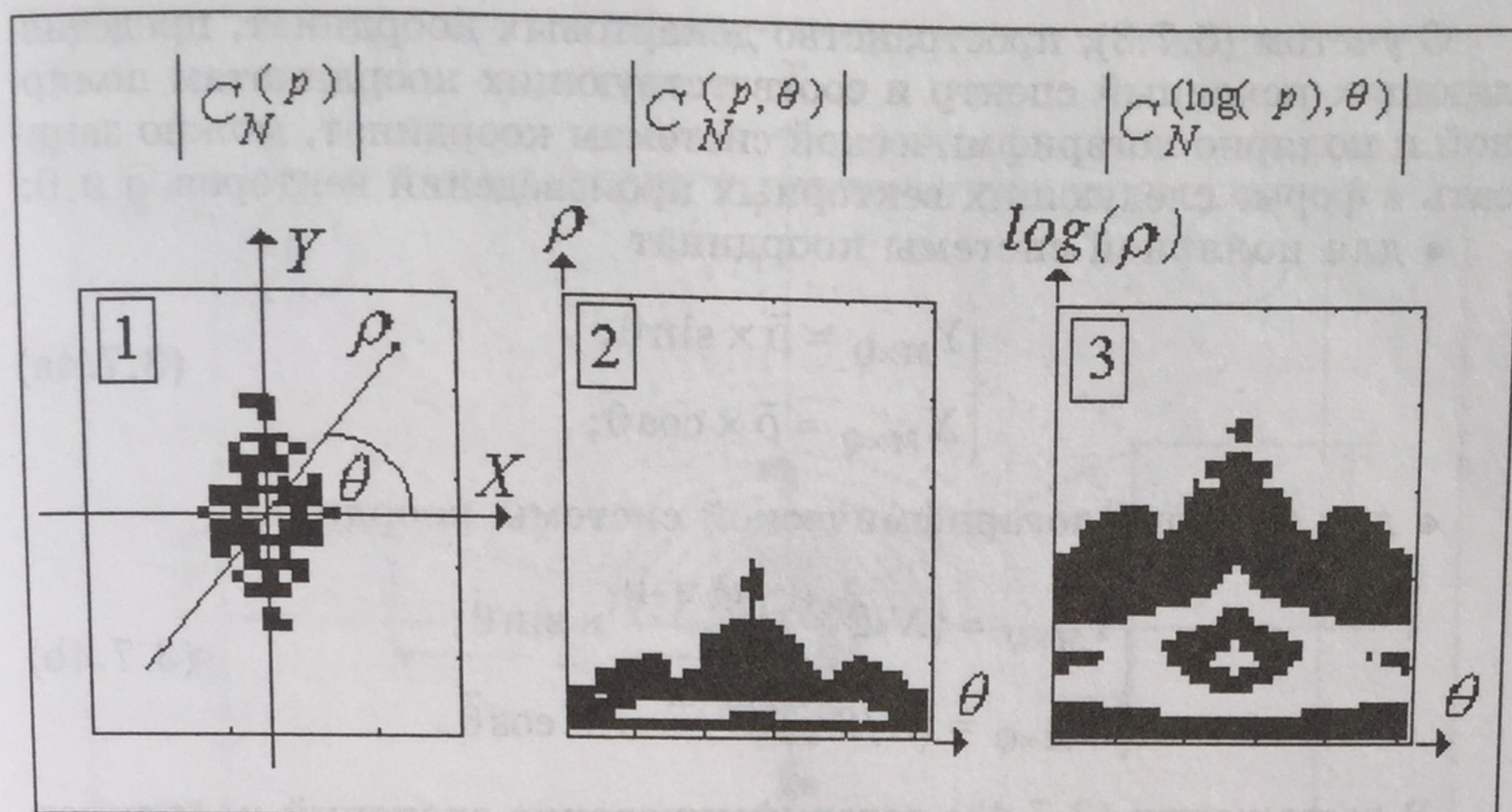


Рис. 3.7.1. Три формы представления модуля спектра преобразования Фурье

Примеры общего вида упомянутых выше спектров показаны на рис. 3.7.1.

В отличие от целочисленных координат декартовой системы, в которой перечисляются номера строк и столбцов матриц спектров, координаты в полярных системах координат не являются целыми числами. Поэтому преобразование Фурье—Меллина заканчивается процедурой интерполяции спектра (из декартовой системы координат) по вычисленным координатам полярной (полярно-логарифмической) системы.

**Преобразования вида**  $C_N^{(\Leftrightarrow)} \rightarrow C_N^{(\rho)} \rightarrow C_{M \times Q}^{(\rho, \theta)} \rightarrow C_{M \times Q}^{(\log(\rho), \theta)}$ .

Для изложения дальнейшего материала введем некоторые обозначения, связанные с представлением спектров в различных системах координат [Кис97а, Кис98с].

Итак, пусть

$$\left\{ \begin{array}{l} \bar{\rho} = \frac{N/2-1}{M-1} \mathbf{R} = \frac{N/2-1}{M-1} [0, 1, 2, \dots, M-1]^T; \\ \bar{\theta} = [\theta_0, \theta_1, \dots, \theta_{Q-1}] = \frac{\pi}{Q} [0, 1, 2, \dots, Q-1]; \\ c^{(\Leftrightarrow)}(\omega_y, \omega_x) = c^{(\Leftrightarrow)}(\rho \sin \theta, \rho \cos \theta), \end{array} \right. \quad (3.7.3)$$

где  $N$  — порядок матрицы, представляющей исходный спектр;  $M$  — число дискрет полярного радиуса;  $Q$  — число полярных углов на интервале  $0 - \pi$ ;  $\bar{\rho}$  — вектор значений полярного радиуса;  $\bar{\theta}$  — вектор значений полярного угла.



С учетом (3.7.3), пространство декартовых координат, представляющих исходный спектр и соответствующих координатам полярной и полярно-логарифмической системам координат, можно записать в форме следующих векторных произведений векторов  $\vec{\rho}$  и  $\vec{\theta}$ :

- для полярной системы координат

$$\begin{cases} Y_{M \times Q} = \vec{\rho} \times \sin \vec{\theta}; \\ X_{M \times Q} = \vec{\rho} \times \cos \vec{\theta}; \end{cases} \quad (3.7.4a)$$

- для полярно-логарифмической системы координат

$$\begin{cases} Y_{M \times Q} = (N/2 - 1)^{R/(M-1)} \times \sin \vec{\theta}; \\ X_{M \times Q} = (N/2 - 1)^{R/(M-1)} \times \cos \vec{\theta}. \end{cases} \quad (3.7.4b)$$

В выражении (3.7.4b) логарифмирование значений координат в направлении полярного радиуса аппроксимировано степенной функцией [Che94].

Рис. 3.7.2 дает графическое представление координат, определенных в соответствии с (3.7.4) и в обозначениях (3.7.3). В левой стороне рисунка — координаты полярной системы координат, а в правой — координаты полярно-логарифмической системы.

Для получения квадратных растров, заполняющих все спектральное пространство и аналогичных представленным справа на рис. 3.1.6, в выражениях (3.7.4) вместо векторов  $\sin \vec{\theta}$  и  $\cos \vec{\theta}$  используем соответственно векторы:

$$\begin{aligned} v^{(\cos)} &= \left[ \frac{\cos \theta_0}{\max(\sin \theta_0, \cos \theta_0)}, \frac{\cos \theta_1}{\max(\sin \theta_1, \cos \theta_1)}, \dots \right. \\ &\quad \left. \dots, \frac{\cos \theta_{Q-1}}{\max(\sin \theta_{Q-1}, \cos \theta_{Q-1})} \right], \\ v^{(\sin)} &= \left[ \frac{\sin \theta_0}{\max(\sin \theta_0, \cos \theta_0)}, \frac{\sin \theta_1}{\max(\sin \theta_1, \cos \theta_1)}, \dots \right. \\ &\quad \left. \dots, \frac{\sin \theta_{Q-1}}{\max(\sin \theta_{Q-1}, \cos \theta_{Q-1})} \right]. \end{aligned} \quad (3.7.5)$$

В случаях, когда исходное изображение (и соответствующий ему спектр) представлено не квадратной матрицей, а прямоугольной размером  $N_1 \times N_2$  пикселей, вместо выражения (3.7.4b) необходимо записать:



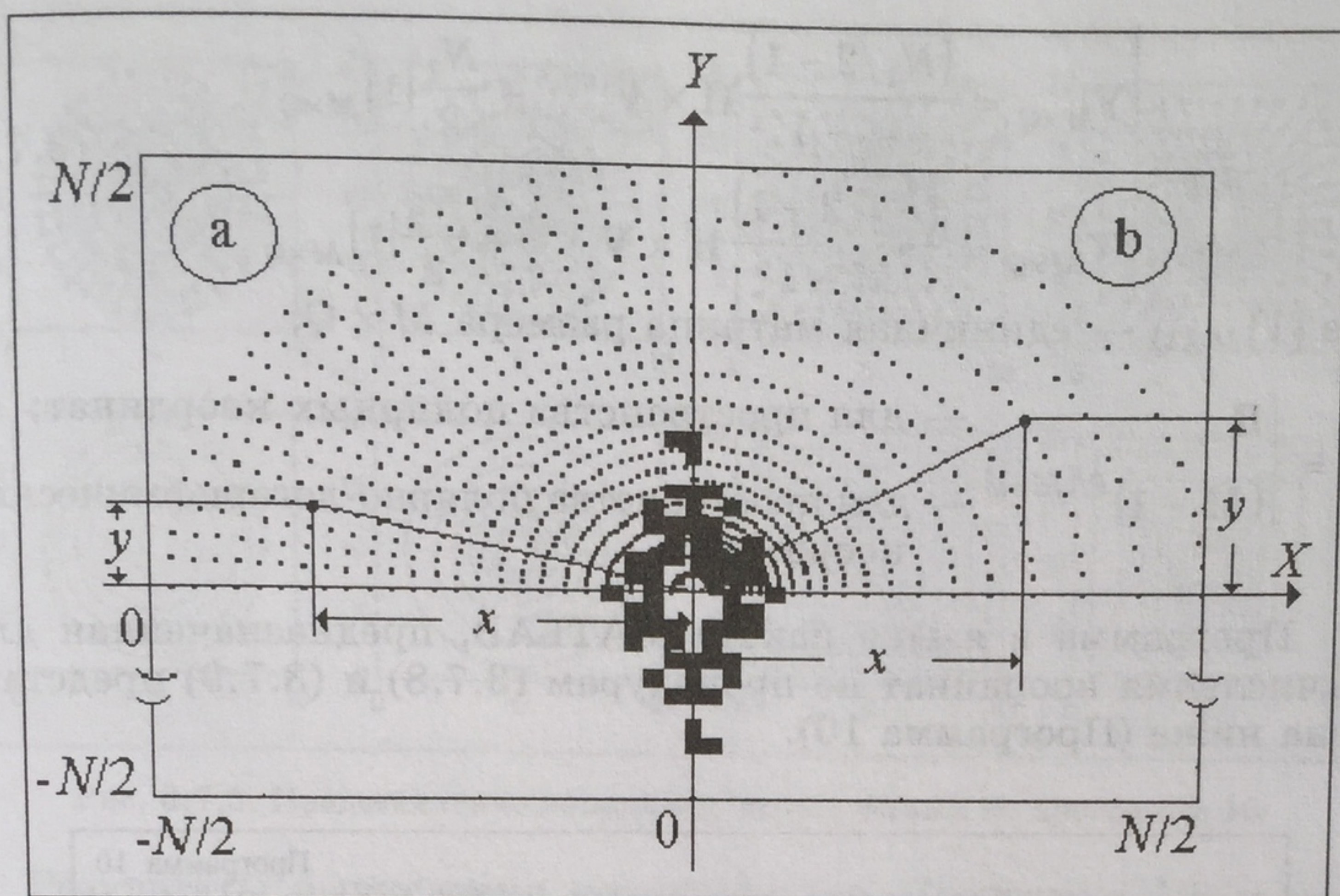


Рис. 3.7.2. Расположение координат полярной и полярно-логарифмической систем на фоне спектра, представленного в декартовой системе координат

для обычных растров

$$\begin{cases} Y_{M \times Q} = (N_1/2 - 1)^{R/(M-1)} \times \sin \bar{\theta}; \\ X_{M \times Q} = (N_2/2 - 1)^{R/(M-1)} \times \cos \bar{\theta}; \end{cases} \quad (3.7.6)$$

для прямоугольных растров

$$\begin{cases} Y_{M \times Q} = (N_1/2 - 1)^{R/(M-1)} \times V^{(\sin)}; \\ X_{M \times Q} = (N_2/2 - 1)^{R/(M-1)} \times V^{(\cos)}. \end{cases} \quad (3.7.7)$$

В практике задач распознавания спектры Фурье и Фурье—Меллина описываются, как правило, размерами представляющих их матриц. Например, исходная матрица (исходного изображения или спектра) имеет размер  $N \times N$  пикселей, а преобразованная — размер  $M \times Q$  пикселей и в общем случае  $M \neq Q \neq N$ .

Для случаев упомянутого матричного представления спектров можно по аналогии с (3.7.6) и (3.7.7) записать соответственно:

$$\begin{cases} Y_{M \times Q} = \frac{(N_1/2 - 1)}{M - 1} \mathbf{H} \times \sin \bar{\theta} + \frac{N_1}{2} [1]_{M \times Q}; \\ X_{M \times Q} = \frac{(N_2/2 - 1)}{M - 1} \mathbf{H} \times \cos \bar{\theta} + \frac{N_2}{2} [1]_{M \times Q}; \end{cases} \quad (3.7.8)$$



$$\begin{cases} Y_{M \times Q} = \frac{(N_1/2 - 1)}{M - 1} \mathbf{H} \times \mathbf{V}^{(\sin)} + \frac{N_1}{2} [1]_{M \times Q}; \\ X_{M \times Q} = \frac{(N_2/2 - 1)}{M - 1} \mathbf{H} \times \mathbf{V}^{(\cos)} + \frac{N_2}{2} [1]_{M \times Q}, \end{cases} \quad (3.7.9)$$

где  $[1]_{M \times Q}$  — единичная матрица размера  $M \times Q$ ;

$\mathbf{H} = \begin{cases} \mathbf{R} & \text{— для пространства полярных координат;} \\ (M - 1)^{R/(M-1)} & \text{— для пространства полярно-логарифмических координат.} \end{cases}$

Программа в языке пакета MATLAB, предназначенная для вычисления координат по процедурам (3.7.8) и (3.7.9) представлена ниже (Программа 10).

Программа 10

```
function [X, Y, X1, Y1] = pol_c(N1, N2, M, Q, par);

% Вычисление полярных (полярно-логарифмических) координат.
% N1, N2 — размер матрицы исходного спектра;
% M, Q — размер матрицы для полярных координат;
% par — "0" (для обычных) и "1" (для полярно-логарифмических);
% X, Y — координаты, вычисляемые по процедуре (3.7.8);
% X1, Y1 — координаты, вычисляемые по процедуре (3.7.9);

% Определение полярных углов и радиуса
teta = 0: 2*Q-1;
R = (0:M-1)';
if par == 1 R = (M-1).^(R/(M-1)); end;

N1 = fix(N1/2); N2 = fix(N2/2);

% Вычисление координат по процедуре (3.7.8)
Y = ((N1-1)/(M1))*R*sin(pi*teta/Q) + N1;
X = ((N1-1)/(M1))*R*cos(pi*teta/Q) + N2;

% -----
% Вычисление векторов Vcos и Vsin
Y1 = sin(pi*teta/Q); X1 = cos(pi*teta/Q);
WW = [X1; Y1];
Vcos = X1./max(abs(WW));
Vsin = Y1./max(abs(WW));

% Вычисление координат по процедуре (3.7.9)
Y1 = ((N1-1)/(M-1))*R*Vsin + N1;
X1 = ((N2-1)/(M-1))*R*Vcos + N2;

% Конец программы
```



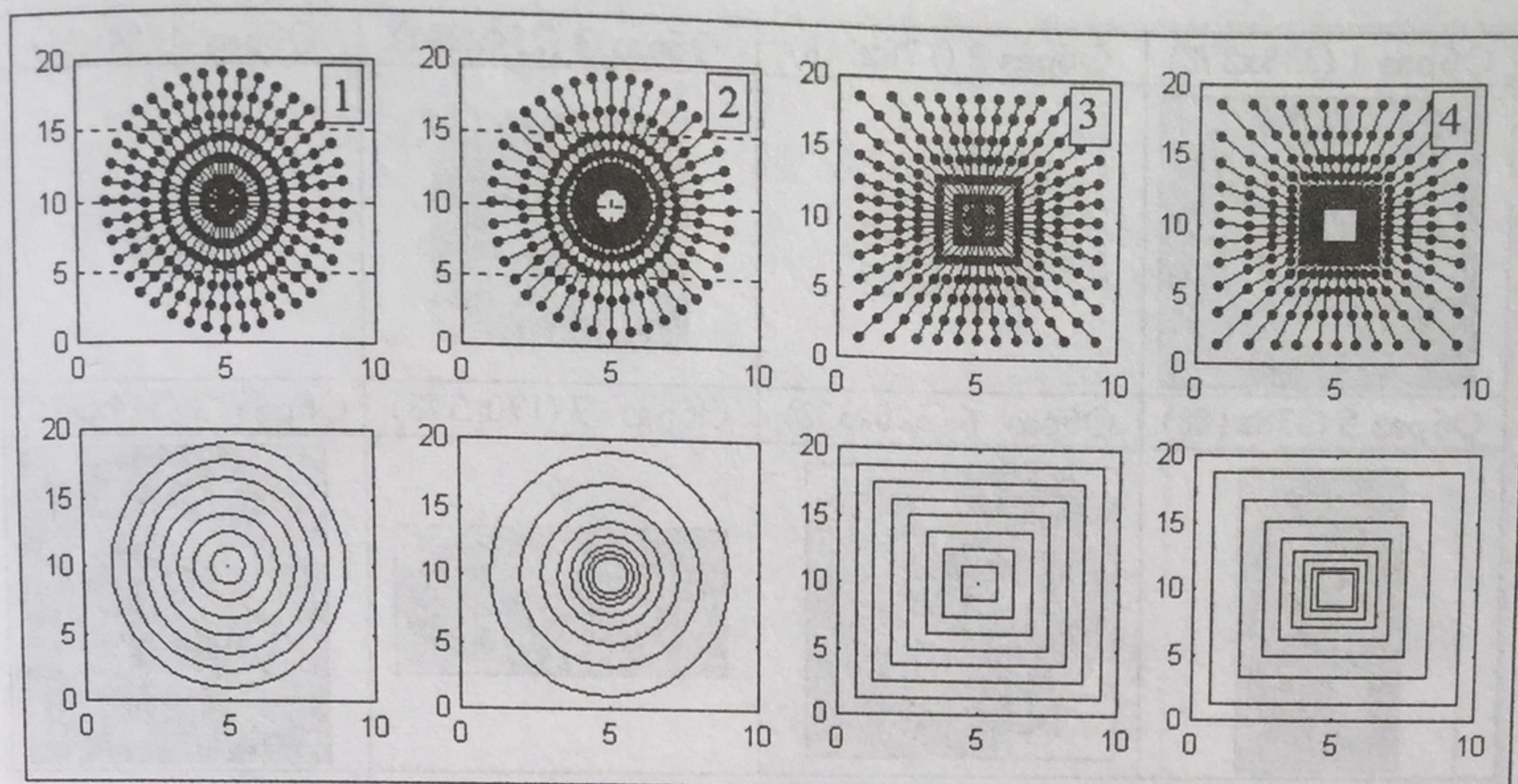


Рис. 3.7.3. Представление координат, вычисленных по программе 10

Результаты вычисления координат (по Программе 10) и различные способы их представления координат показаны на рис. 3.7.3. Во всех случаях координаты определены для следующих значений размеров матриц:  $N_1 = 20$ ,  $N_2 = 10$ ,  $M = 8$ ,  $Q = 20$ . Цифрами 1 и 2 обозначены результаты, полученные по процедуре (3.7.8) для полярной и полярно-логарифмической систем координат соответственно. Различие между этими системами хорошо видно на нижних рисунках, где координаты, относящиеся к одному и тому же радиусу, соединены между собой. Цифрами 3 и 4 обозначены результаты, полученные по процедуре (3.7.9) также для полярной и полярно-логарифмической систем координат соответственно. Нижние рисунки подчеркивают различие в вычисляемых по (3.7.9) полярной и полярно-логарифмической системах координат.

Для получения окончательного результата преобразования Фурье—Меллина необходимо выполнить переинтерполяцию по вычисленным координатам. Если положить, что совокупность координат  $X_0$  и  $Y_0$  представляет некоторую матрицу спектра  $S_{cart}$  в декартовой системе, а совокупность координат  $X$  и  $Y$  (или  $X_1$  и  $Y_1$ ) получена по программе 10 [по процедурам (3.7.8) или (3.7.9)], то спектр  $S_{pol}$  Фурье—Меллина можно вычислить, используя, например, процедуру «interp2» двумерной интерполяции пакета MATLAB [Пот99, Руд00]:

$$S_{pol} = \text{interp2}(X_0, Y_0, S_{cart}, X, Y, 'method'). \quad (3.7.10)$$

Интерполированный по полярно-логарифмическим координатам и развернутый спектр представляет собой окончательный результат преобразования Фурье—Меллина, который называют также инвариантом Фурье—Меллина (ФМИ [Che94]).

Рассмотрим теперь пример использования инвариантов Фурье—Меллина для распознавания изображений, подвергнутых сильным изменениям в масштабе. На рис. 3.7.4 приведены во-



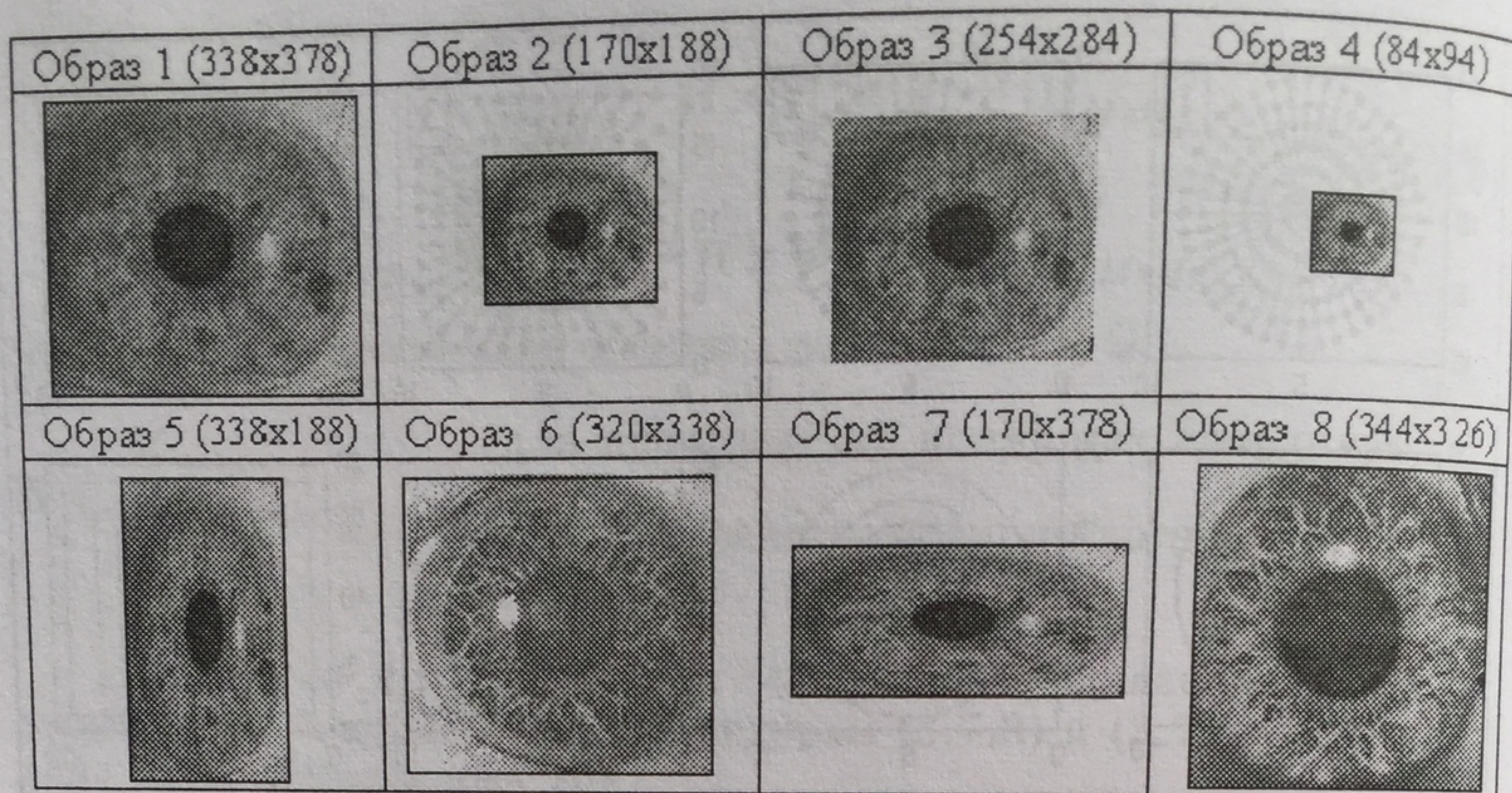


Рис. 3.7.4. Тестовые изображения глаз

семь изображений глаз (размеры каждого изображения указаны в скобках).

Шесть из них (1, 2, 3, 4, 5 и 7) принадлежат изображению одного и того же глаза. Наибольший размер здесь имеет образ 1 ( $338 \times 378$  пикселей), а наименьший — образ 4 ( $84 \times 94$  пикселя). Образ 1 базовый. Образы 2 — 5 и 7 уменьшены в размерах по сравнению с базовым, причем масштаб образов 3 и 4 по обеим осям изменялся одинаково, а для образов 2, 5 и 7 — неодинаково (непропорционально). В свою очередь, образы 1, 6 и 7 представляют изображения глаз различных людей.

В выполненном эксперименте сравнивались между собой все образы и оценивалась мера их подобия. Использование пути, связанного с выравниванием всех размеров к одному из них, не рассматривалось, поскольку и уменьшение размера изображения и увеличение его размера приводит к потере качества цифрового изображения. Поэтому для элиминирования влияния различия в размерах  $N_1$  и  $N_2$  каждый образ подвергается преобразованию Фурье—Меллина с различными входными параметрами  $N_1$  и  $N_2$  и одинаковыми выходными параметрами  $M$  и  $Q$  [см. выражение (3.7.8) — (3.7.9)].

На рис. 3.7.5 представлены результаты для этого случая сканирования исходного спектра в полярно-логарифмической системе координат. Цифры от 1 до 8 соответствуют номерам образов на рис. 3.7.4.

Заканчивая поставленный эксперимент, сравним между собой вычисленные по (3.7.10) инварианты Фурье—Меллина. Для этого используем двумерную фазовую корреляцию, описание которой было дано в разд. 3.5. Полученные при этом результаты корреляции представлены в табл. 3.7.1.

Как видно из таблицы, образы 1–5 и 7 соответствуют друг другу, поскольку коэффициент взаимной корреляции между ними не ниже величины 0,88. В то же время образы первой группы



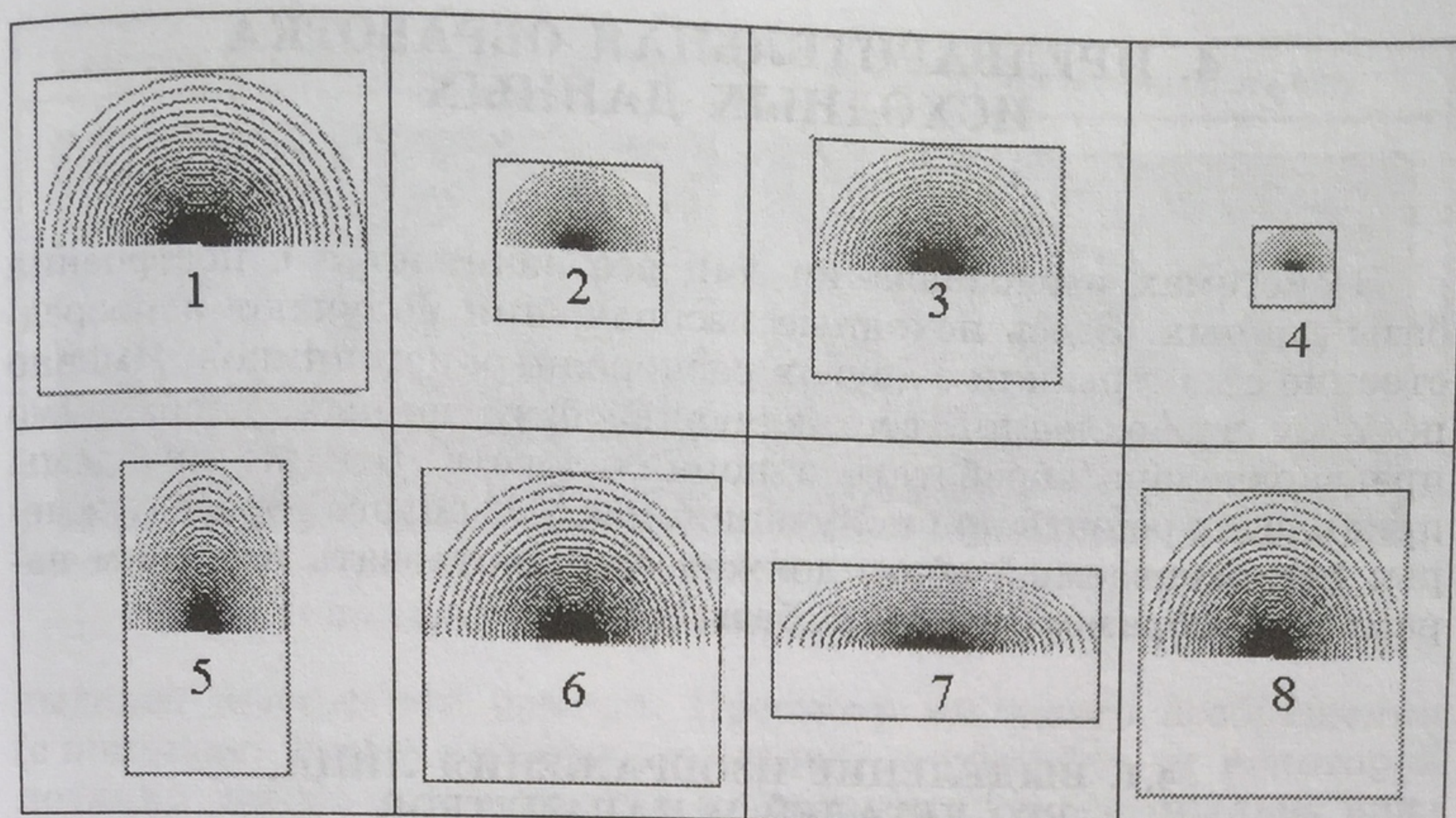


Рис. 3.7.5. Изменение способа сканирования спектров в преобразовании Фурье—Меллина в зависимости от размеров исходных изображений

1–5 и 7 не соответствуют образам 6 и 8, поскольку коэффициент корреляции между ними не превосходит 0,14.

Таким образом, мы показали, что преобразование Фурье—Меллина [Che94] элиминирует масштабные изменения изображений. Элиминирование поворотов образов на изображениях основано на преобразованиях Фурье-спектров к полярной системе координат. Циклические сдвиги изображений элиминируются благодаря применению в преобразованиях абсолютных значений Фурье-спектров.

Использование инвариантов Фурье—Меллина в биометрических технологиях идентификации человека на основе радужной оболочки глаза представлено в работах [Bor97, Sol97]. Применение инвариантов Фурье—Меллина в задачах распознавания лиц показано, например, в работах [Ben98].

Таблица 3.7.1

Результаты фазовой корреляции

| Образ | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1     | 1,0000 | 0,9761 | 0,9906 | 0,8896 | 0,9841 | 0,0882 | 0,9874 | 0,0982 |
| 2     | *      | 1,0000 | 0,9712 | 0,9028 | 0,9892 | 0,0879 | 0,9818 | 0,0938 |
| 3     | *      | *      | 1,0000 | 0,8736 | 0,9708 | 0,0908 | 0,9895 | 0,0949 |
| 4     | *      | *      | *      | 1,0000 | 0,9133 | 0,0894 | 0,8735 | 0,1146 |
| 5     | *      | *      | *      | *      | 1,0000 | 0,0867 | 0,9732 | 0,1002 |
| 6     | *      | *      | *      | *      | *      | 1,0000 | 0,0874 | 0,1405 |
| 7     | *      | *      | *      | *      | *      | *      | 1,0000 | 0,0960 |
| 8     | *      | *      | *      | *      | *      | *      | *      | 1,0000 |



## 4. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИСХОДНЫХ ДАННЫХ

В системах распознавания лиц все начинается с построения базы данных. Здесь исходные изображения получают непосредственно с камеры или с других разнородных источников. Именно поэтому изображения, составляющие базу данных, необходимо предварительно обработать и нормализовать. Эти же проблемы приходится решать при получении каждого нового образа с камеры: распознаваемый образ должен соответствовать по своим параметрам образам исходной базы данных.

### 4.1. ВЫДЕЛЕНИЕ ИЗОБРАЖЕНИЯ ЛИЦА, ЕГО ДЕТАЛЕЙ И ПАРАМЕТРОВ

Наиболее трудной задачей среди перечисленных выше является выделение собственно лица из исходного (заданного) изображения, а также его основных частей —  $2/3$  нижней части лица, глаз, носа и т. д. Для решения этих задач используются два известных подхода [Прэ82, Bry99, Pen00].

Первый [Прэ82] основан на просмотре всего исходного изображения локальной маской и сопоставлении образа в поле этой маски с некоторым эталоном (см. разд. 2.4). Эталон является изображением заданного размера, в котором человеческое лицо или его часть (с фотографии, например) расположено в центре, а фон в окружении лица соответствует фону исходного изображения. При этом на эталоне должно быть представлено практически фронтальное изображение лица человека. В качестве меры различия берется среднеквадратическая ошибка или нормированная корреляция между образом в поле маски и эталоном. Минимум среднеквадратической ошибки (или максимум взаимной нормированной корреляции), полученный в результате всего просмотра, будет свидетельствовать о наилучшем соответствии изображения в поле маски — эталону.

Другой подход [Bry99, Pen00] использует информацию о градиентах яркости исходного изображения и процедуру вычисления центра лица, относительно которого затем выделяется все лицо.

#### 4.1.1. ВЫДЕЛЕНИЕ ЛИЦА И ЕГО ЧАСТЕЙ

Пример использования первого подхода для выделения лица из фотографии показан на рис. 4.1.1. В качестве эталона выбрано лицо человека, стоящего в центре левой фотографии, а сам эталон



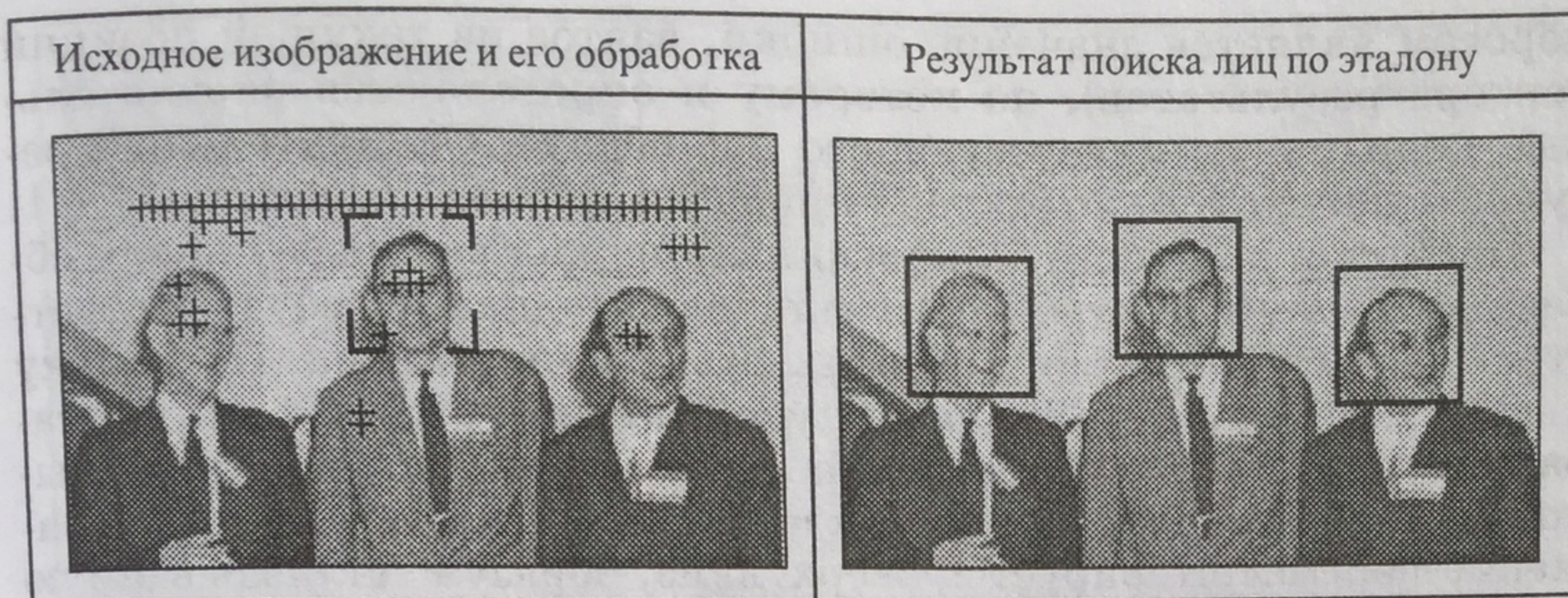


Рис. 4.1.1. Выделение лиц из фотографии. Фото из пакета МАТЛАБ

выделен квадратной рамкой. Просмотр исходного изображения (с помощью маски размером с эталон) начинается от некоторой позиции левого верхнего угла и продолжается вертикально вниз до момента, пока маска находится в поле исходного изображения. На рисунке знаками «+», выстроенными в линейку, обозначены позиции начала поиска.

На каждом шаге просмотра оценивается среднеквадратическая ошибка между образом в поле маски и эталоном. Если на текущем шаге просмотра ошибка меньше некоторого заданного порога и меньше, чем вычисленная на предыдущем шаге, то она записывается в «вектор ошибок», а координаты центра маски запоминаются. На рис. 4.1.1 в поле лица каждого человека видны знаки «+», определяющие координаты центров маски, для которых вычисленная среднеквадратическая ошибка отвечала приведенным выше условиям. По окончании первого вертикального просмотра в векторе ошибок отыскивается «глобальный минимум», а по нему — соответствующие координаты центра маски, которые определяют наилучшее совмещение (привязку) изображения в поле маски с эталоном. Полученные при этом три параметра — ошибка, координата  $X$  (столбец исходного изображения) и координата  $Y$  (строка исходного изображения) записываются в некоторый вектор результатов.

Далее просмотр продолжается с новой верхней позиции, отстоящей от самой первой позиции на некоторую дельту  $\Delta$ . И так происходит просмотр (сканирование) всего исходного изображения, а в векторе результатов накапливается информация о тех участках исходного изображения, которые соответствуют эталону.

Для ускорения сканирования всего исходного изображения переход на новую позицию маски осуществляется скачкообразно — с некоторой  $\Delta$ , обычно большей, чем один пиксель. Поэтому по окончании всего просмотра исходного изображения выполняется второй этап — локальная привязка. Она реализуется относительно параметров, записанных в векторе результатов, с шагом сканирования  $\pm 1$  пиксель и только в пределах  $\Delta$ . В процессе локальной привязки также ищется минимум ошибки (при этом



порогом является значение ошибки, взятое из текущей позиции вектора результатов), по которому и определяются окончательные координаты интересующего нас образа. Окончательный результат поиска лиц по эталону приведен в правой части рис. 4.1.1.

Несмотря на простоту, описанный подход реализует поиск образа по эталону быстро и результативно, однако требует соответствия размеров искомого образа с размерами эталона. Поскольку в практике обработки заданных изображений априори нельзя гарантировать такого соответствия, описанный выше процесс выполняется несколько раз, в каждом из которых исходное изображение перемасштабируется вверх/вниз, образуя так называемую пирамиду изображений [Row98]. Для каждого изображения пирамиды отыскиваются координаты привязки, и те из них считаются актуальными, для которых есть максимальное число совпадений этих координат на соответствующих локальных участках. На этом же этапе определяются актуальные размеры исходного изображения, привязанного к размеру эталона.

На рис. 4.1.2 представлены некоторые варианты эталонов, используемых в системах распознавания лиц, которые мы также будем в дальнейшем использовать. Исходным является эталон, отмеченный цифрой 1, т. е. фронтальное изображение лица человека. Его размер  $112 \times 92$  пикселя. Все последующие эталоны получены из эталона 1. Номером 2 отмечен эталон, охватывающий наиболее информативную (индивидуальную) часть лица, а номером 3 — эталон, представляющий наиболее неизменяемую часть лица. Эталоны 4, 5 и 6 определяют различные детали лица.

Покажем теперь, что метод сравнения с эталоном позволяет выделить интересующий нас образ даже в тех случаях, когда размеры и пункт наблюдения в исходном изображении не совсем точно соответствуют эталону.

В качестве исходных изображений выберем по одному образу из первых 10 классов базы [Orl\*\*], а в качестве эталона — эталон 2 (см. рис. 4.1.2). Процесс выделения искомого образа из исход-

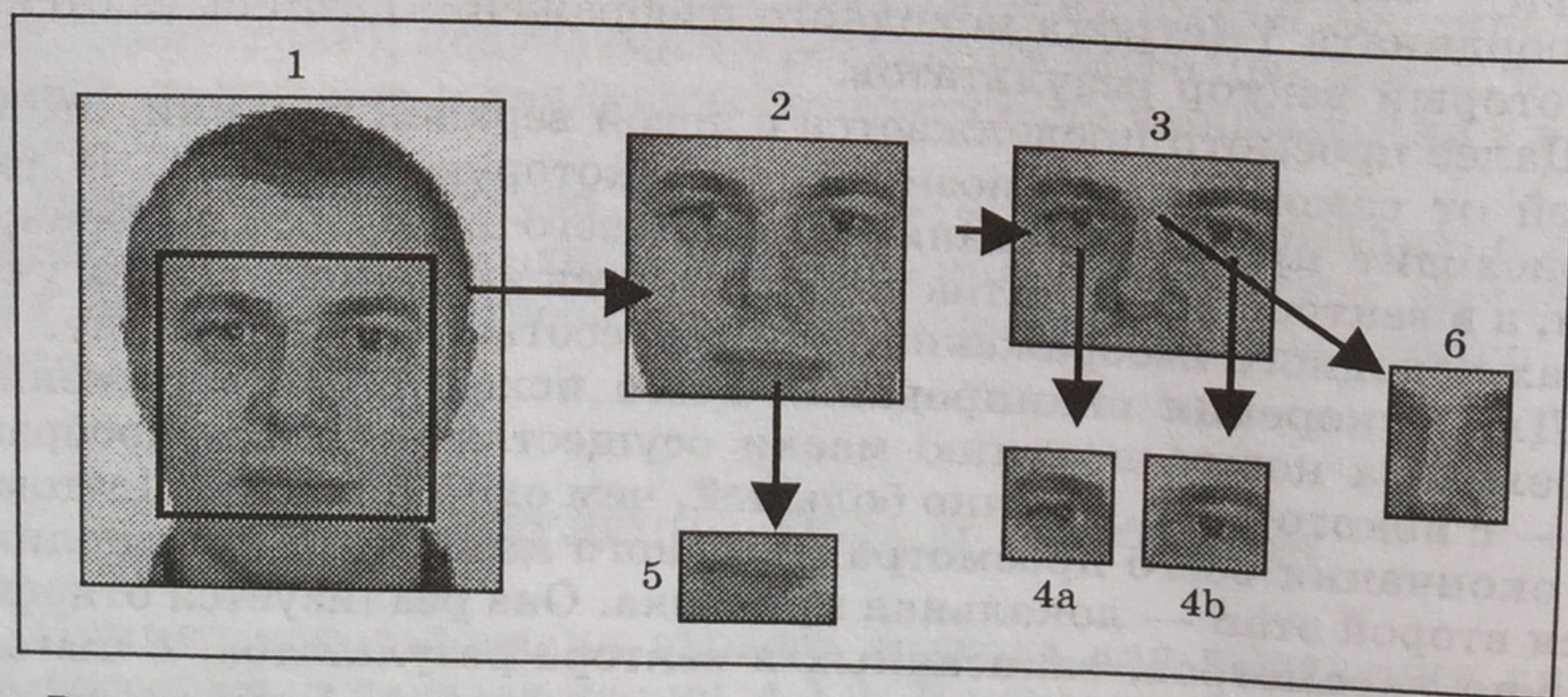


Рис. 4.1.2. Варианты эталонов, используемых в системах распознавания лиц



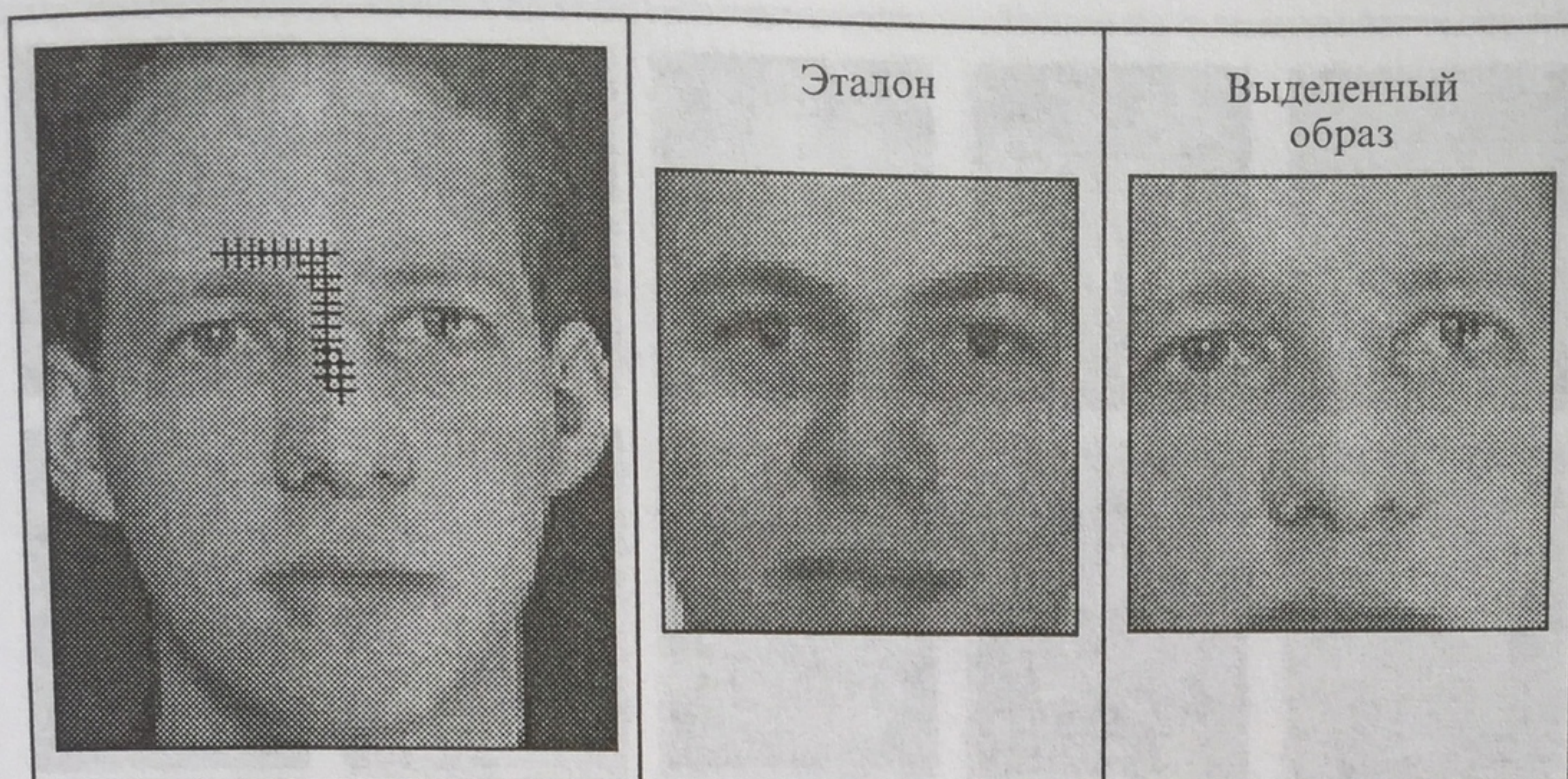


Рис. 4.1.3. Процесс и результат выделения искомого образа по заданному эталону

ного изображения (первый образ из первого класса базы [Orl\*\*]) представлен на рис. 4.1.3 с левой стороны, где знаками «+» отмечены координаты центров сканирующей маски.

С правой стороны представлен выделенный образ. Даже «на глаз» видно несоответствие масштабов эталона и выделенного образа. Разными также являются и пункты наблюдения собственно лица для эталона и выделенного образа. Однако и при этом хорошо видна привязка по уровню глаз.

Эту оценку дополнит результат выделения интересующего нас образа из исходных изображений всех 10 классов базы [Orl\*\*], который представлен на рис. 4.1.4а.

А теперь в качестве эталона возьмем эталон 3 (см. рис. 4.1.2), а в качестве исходных изображений — все 10 образов первого класса базы [Orl\*\*].

Результат выделения образов, соответствующих эталону 3, представлен на рисунке 4.1.4б. Можно оценить этот результат как превосходный, поскольку практически ни для одного из исходных изображений не было выполнено условие полного соответствия с эталоном.

Из приведенных примеров следует, что в данном подходе, использующим процедуру сопоставления с эталоном, не требуется точного соответствия параметрам эталона: размеры выделяемого образа и эталона могут отличаться в пределах 5 – 10 %, не требуется и строго фронтального образа на эталоне и исходном изображении и т. д.

Выше мы отмечали, что фон в окружении лица на изображении эталона должен соответствовать фону исходного изображения с камеры. Во-первых, на практике при использовании систем распознавания такое условие не всегда выполнимо. Во-вторых, у людей разных национальностей уровень ясности лиц также будет разный: сравним, например, уровень ясности лиц у европейца и азиата, у европейца и афроамериканца и т. д.



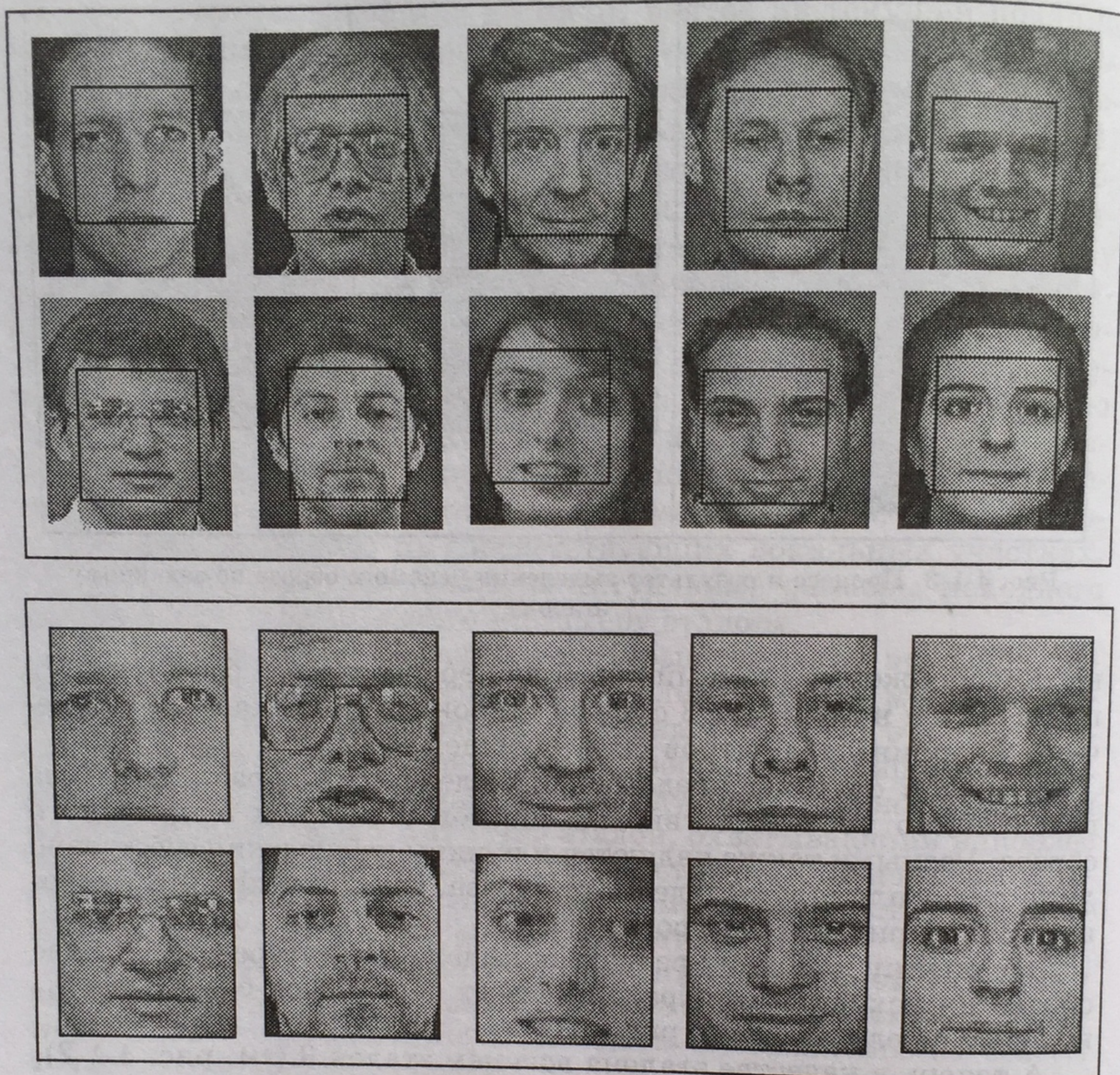


Рис. 4.1.4а. Результат выделения требуемого образа из различных исходных изображений

Можно было бы использовать несколько эталонов, учитывающих конкретные условия применения системы распознавания, однако такой подход приводит к многократному (по числу эталонов) использованию процедуры сопоставления изображения в поле маски с эталоном и усложняет принятие окончательного решения: при разных (априори неизвестных сочетаниях фона и ясности образа в поле рамки исходного изображения) мера подобия/различия с эталоном может быть непредсказуемой и использованный минимаксный критерий может не всегда достоверно сработать.

Рассмотрим новое и очевидное решение. Вместо обычного изображения эталона будем всегда использовать матрицу его градиентов яркости, а перед процедурой выделения лица (или его деталей) исходное изображение также представим в форме его градиентов яркости. При этом в качестве меры различия двух матриц градиентов (одна соответствует эталону, другая — образу в поле рамки



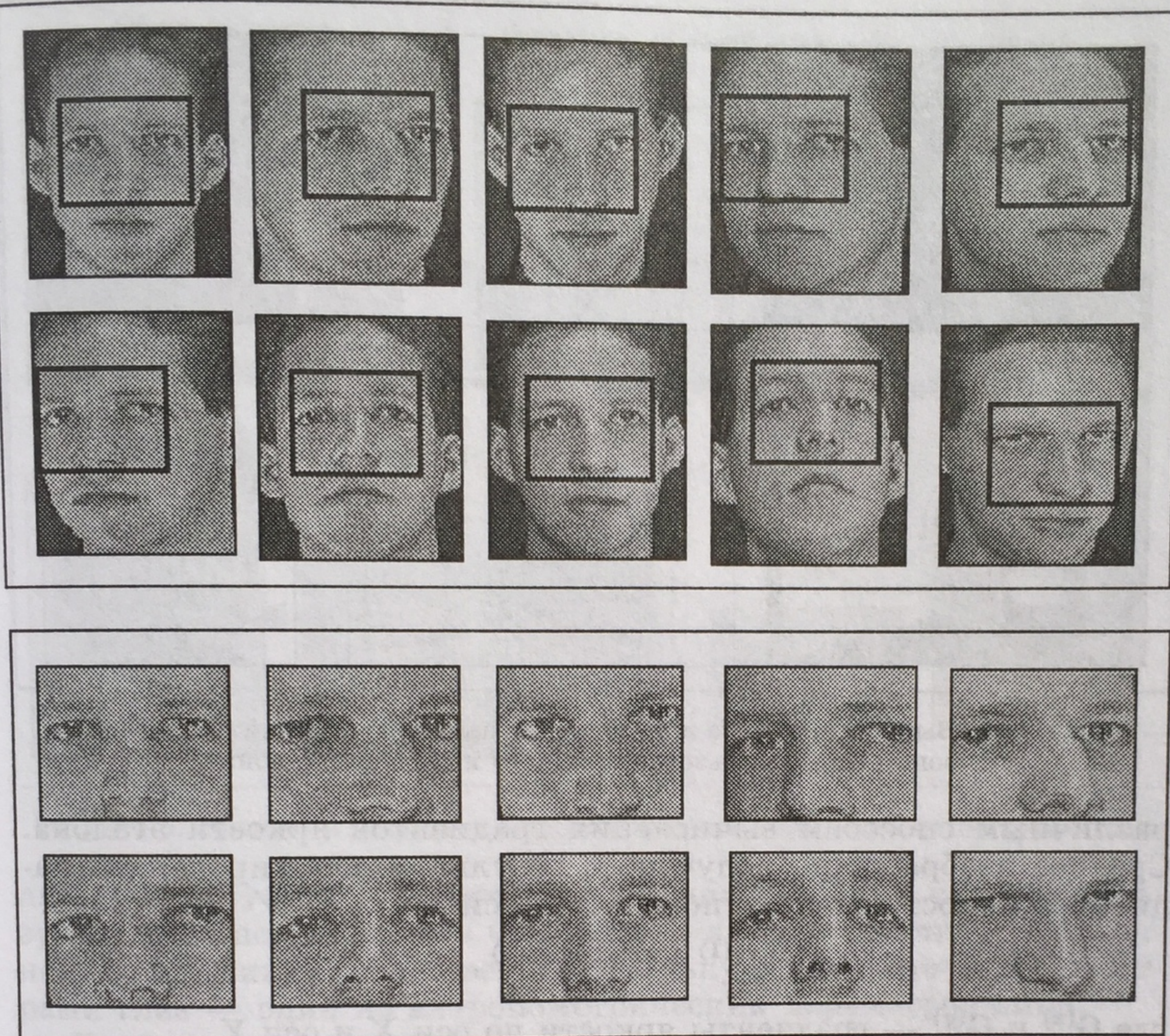


Рис. 4.1.4b. Результат выделения требуемого образа из исходных изображений одного и того же класса образов базы [Orl\*\*]

на исходном изображении) необходимо использовать или метрику  $L_1$  (сумму модулей разностей между ними), или метрику  $L_0$ .

На рис. 4.1.5а приведено изображение эталона 1 (представленного ранее на рис. 4.1.2) и два изображения, соответствующие

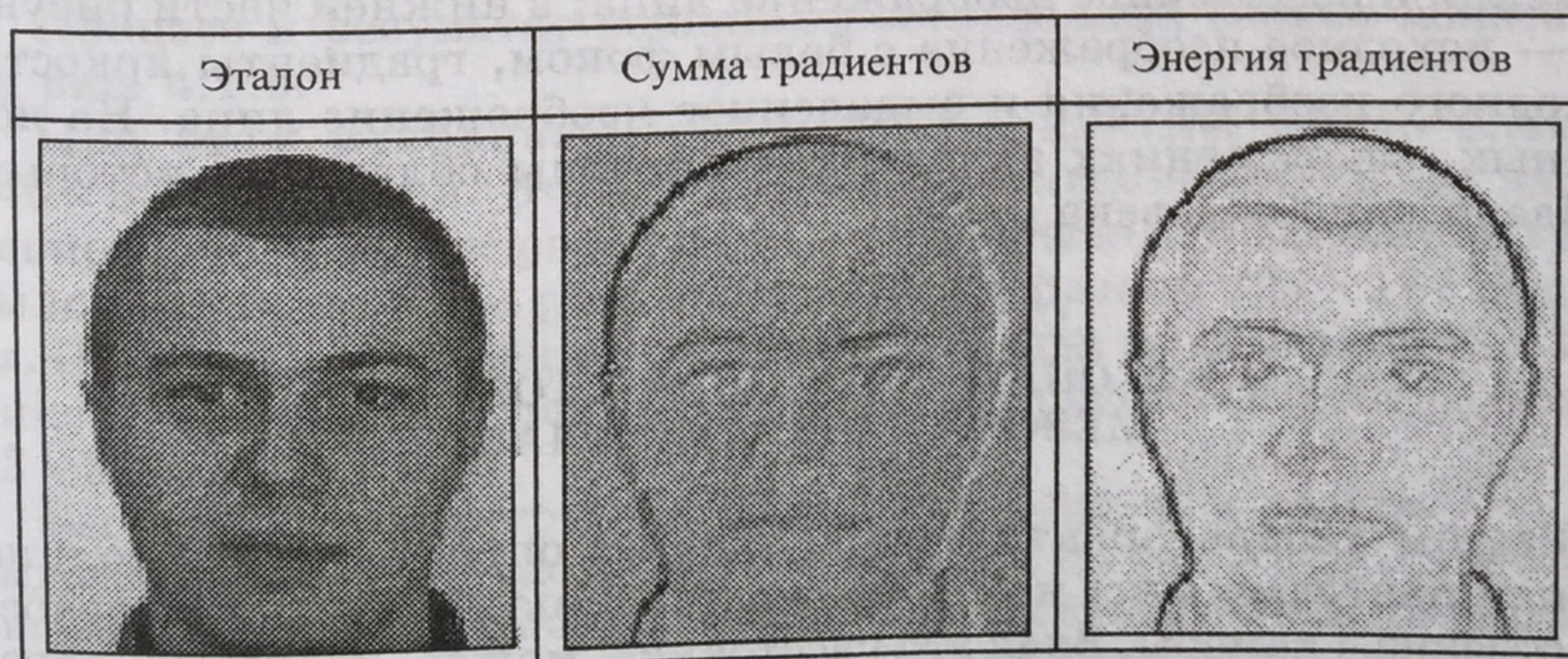


Рис. 4.1.5а. Эталон и представление его в форме градиентов



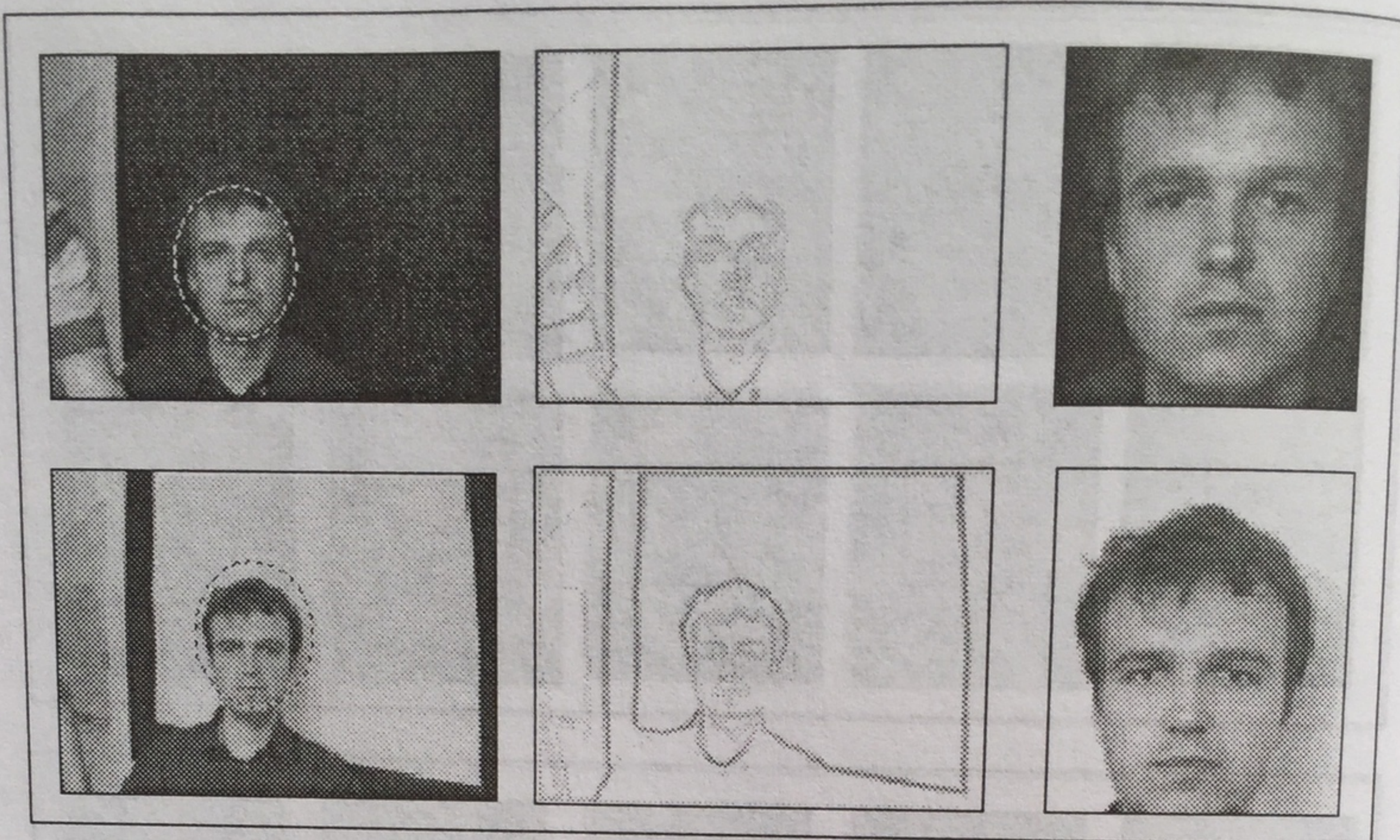


Рис. 4.1.5b. Выделение образа лица из исходных изображений с различным фоном при использовании одного и того же эталона

различным способам вычисления градиентов яркости эталона. Среднее изображение получено в результате суммирования градиентов яркости эталона по оси  $X$  и оси  $Y$ :

$$O^{(1)} = G^{(x)} + G^{(y)},$$

где  $G^{(x)}$  и  $G^{(y)}$  — градиенты яркости по оси  $X$  и оси  $Y$ .

Правое изображение вычислено как корень квадратный из суммарной энергии градиентов:

$$O^{(2)} = \sqrt{G^{(x)}G^{(x)} + G^{(y)}G^{(y)}}.$$

На рис. 4.1.5b приведены: в верхней части рисунка исходное изображение с черным фоном, градиенты яркости исходного изображения и выделенное изображение лица; в нижней части рисунка — исходное изображение с белым фоном, градиенты яркости исходного изображения и выделенное изображение лица. На исходных изображениях эллипсами очерчены области, в которых найдено лицо человека.

#### 4.1.2. ОПРЕДЕЛЕНИЕ РАССТОЯНИЯ МЕЖДУ ЦЕНТРАМИ ГЛАЗ

Решим теперь практическую задачу: определим расстояние между центрами глаз на изображении лица, полученном непосредственно с камеры. Зная это расстояние, можно, например, определить ось симметрии лица, а также перемасштабировать образ



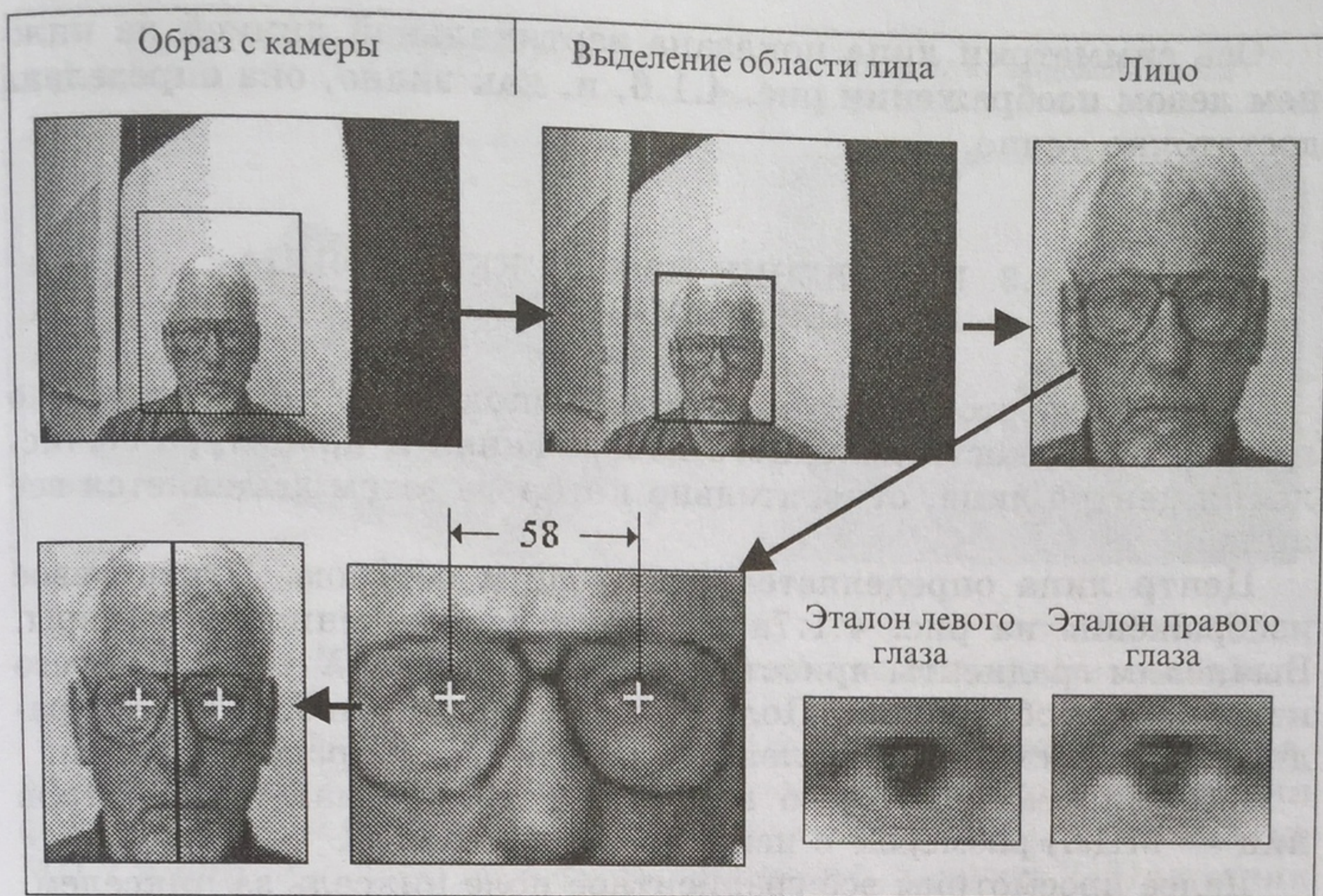


Рис. 4.1.6. Процесс и результат определения расстояния между центрами глаз

лица до требуемых размеров. Последняя операция часто используется при распознавании человека по лицу, в том числе для точной идентификации человека, поскольку расстояние между центрами глаз — один из антропометрических параметров лица.

Для решения поставленной задачи в автоматическом режиме сначала выделим собственно изображение лица из исходного изображения с камеры, используя эталон 1, затем с использованием эталона 3 выделим часть лица и, наконец, используя эталоны 4, найдем центры регионов глаз и оценим расстояние между ними. Середина между центрами глаз — есть ось симметрии образа лица, если, правда, образ лица имеет строго фронтальный вид. В данном случае мы используем именно фронтальный вид лица.

Процесс и результат решения поставленной задачи показаны на рис. 4.1.6.

При этом исходное изображение с камеры имело размер  $240 \times 320$  пикселей. Эталон 1 (и соответственно ему образ выделенного лица) имеют размеры  $112 \times 92$  пикселя. Выделенное лицо было увеличено в два раза и стало иметь размер  $224 \times 184$  пикселей. Эталон для выделения средней части лица имел размер  $88 \times 120$  пикселей. Эталоны глаз (приведены на рисунке) имели размеры  $22 \times 33$  пикселя.

Вычисленные центры регионов глаз показаны знаком «+» на выделенной средней части лица. Несмотря на то, что на лице имелись очки, центры глаз определены достаточно точно. Затем эти центры были «перенесены» на образ целого лица. При этом вычисленное расстояние между центрами глаз составило 58 пикселей.



Ось симметрии лица показана вертикальной линией на нижнем левом изображении рис. 4.1.6, и, как видно, она определена достаточно точно.

### 4.1.3. ВЫДЕЛЕНИЕ ИЗОБРАЖЕНИЯ ЛИЦА ГРАДИЕНТНЫМИ МЕТОДАМИ

Напомним, что при этом подходе используется информация о градиентах яркости исходного изображения и процедура вычисления центра лица, относительно которого затем выделяется все лицо.

Центр лица определяется следующим образом. Пусть левое изображение на рис. 4.1.7а является изображением с камеры. Вычислим градиенты яркости (отдельно по оси  $X$  и оси  $Y$ ) всего исходного изображения. Поле совместных по оси  $X$  и оси  $Y$  градиентов яркости представлено на рис. 4.1.7а с правой стороны.

Сформируем некоторую нулевую матрицу (матрицу центров лиц — МЦЛ) размером с исходное изображение.

Далее просмотрим все градиентное поле пиксель за пикселем. Если в текущем пикселе  $(i, j)$  градиент превышает некоторый порог, то в предположении, что лицо может быть описано кругом радиуса  $R$ , вычисляются новые координаты  $(m, n)$ , отстоящие от координат  $(i, j)$  на этот радиус и по адресам координат  $(m, n)$  в МЦЛ добавляется «1».

Если следовать по градиентам овала лица, то наибольшее число единичек в МЦЛ будет записано в координатах, соответствующих центру лица, и близко к нему, в то время как единички всех симметричных новых координат будут распределены равномерно за регионом лица, повторяя его овал.

На рис. 4.1.7b в левой части представлен вид МЦЛ в 3D-форме. Цифрами 1 и 2 обозначены пики, находящиеся на месте центров лиц.

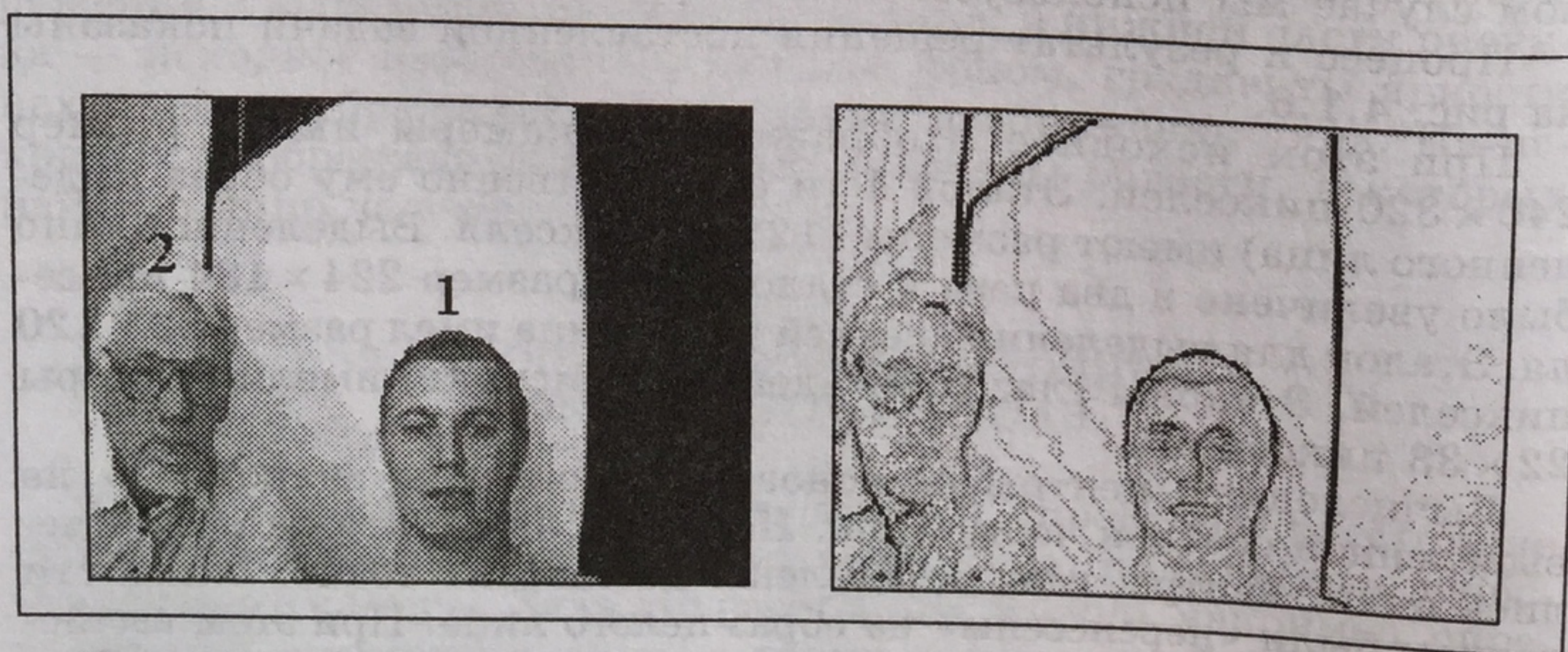


Рис. 4.1.7а. Исходное изображение и его представление в поле градиентов



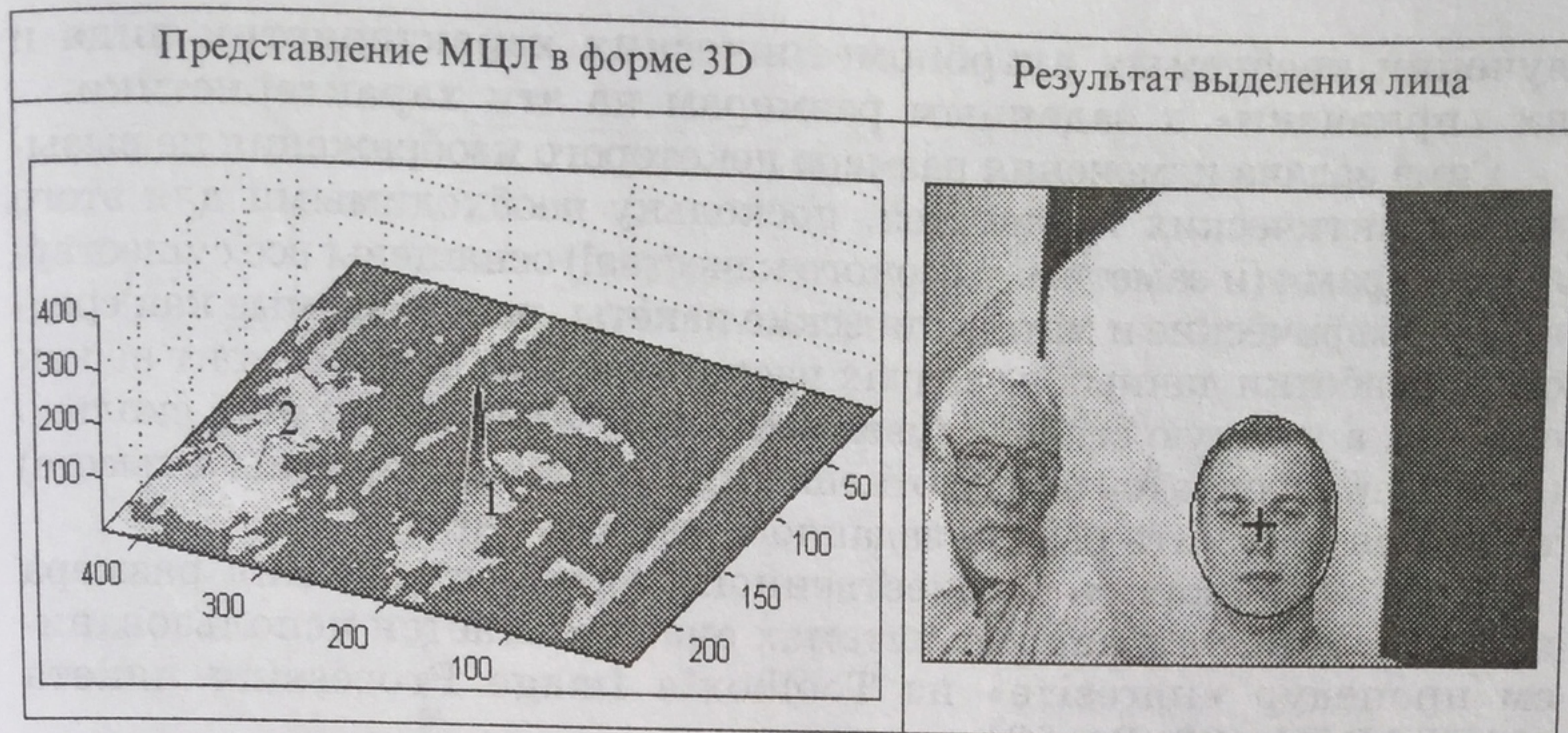


Рис. 4.1.7b. Результат выделения региона лица на изображении с камеры

Поскольку лицо № 1 было ближе к камере, имело более правильные формы и вокруг него не было значительного изменения фона (такого, который имеется вокруг лица № 2), то пик 1 на МЦЛ имеет большую высоту по сравнению с пиком 2. Применяя далее пороговую селекцию, можно выделить только пик 1. Центр этого пика отмечен на правом изображении рис. 4.1.7b крестиком. Относительно этого центра можно теперь выделить интересующее нас изображение лица.

На МЦЛ можно также заметить и другие пики, овалы и линии, сопоставляя которые с изображением в форме градиентов, несложно определить их принадлежность исходному изображению.

Заметим, что на самом деле овал лица не всегда имеет форму круга. Он традиционно ближе к эллипсу, поэтому перед вычислением градиентов исходное изображение перемасштабируется — увеличивается по ширине на 20–25 %. В конце вычислений выполняется обратная операция с тем, чтобы центр выделяемого лица соответствовал масштабу реальной сцены, полученной с камеры.

Несмотря на относительную сложность описанного подхода, его применяют для выделения центров лиц на групповых фотографиях людей, а также для выделения центров деталей лиц — глаз, носа, рта. Примеры использования этого подхода для выделения деталей лиц можно найти в обзоре [Bry99, Pen00].

## 4.2. СОГЛАСОВАНИЕ РАЗМЕРОВ ИЗОБРАЖЕНИЙ

При практическом решении задач распознавания образов лиц мы постоянно сталкиваемся с необходимостью изменения размеров изображений. Эта же необходимость возникала (см. гл. 5 и начало гл. 6) при согласовании размера контрольного или выделенного изображения до требуемого; при построении пирамиды изображений из единственного исходного изображения; при по-



лучении требуемых антропометрических характеристик лица и их «привязки» к заданным размерам на эти характеристики.

Сама задача изменения размера некоторого изображения не вызывает практических трудностей, поскольку необходимыми для этого процедурами (и заметим, высокого качества!) оснащены все существующие графические и математические пакеты, используемые как средства обработки данных в задачах распознавания. Вопрос стоит не как сделать, а в какую сторону (увеличить или уменьшить) и на сколько (пикселей, процентов или по отношению, например, ширины к высоте) требуется изменить размер заданного изображения.

Если же говорить о качественной стороне изменения размера изображений, то в наших системах она достигается использованием процедур «imresize» из Toolbox'a Image Processing пакета MATLAB [Пот99, Руд00].

На рис. 4.2.1 приведена структурная схема простого алгоритма согласования размера выделяемого образа с использованием параметра  $R$  — расстояния между центрами глаз.

Исходным здесь является изображение с камеры, а выделяемый образ — средняя часть лица. Необходимо выбрать такой размер выделяемого образа, для которого расстояние между центрами глаз будет равно величине  $S$  с ошибкой не большей некоторой  $\delta$ .

Если изображение с камеры имеет размер  $240 \times 320$  пикселей, а изображение всего лица размещено в рамке  $112 \times 92$  пикселя, то параметр  $\delta$  обычно выбирается равным 1.

Первая оценка расстояния  $R$  между центрами глаз дается в соответствии с методом, представленным выше. На этом же этапе вычисляется положение оси симметрии лица, относительно которого будет формироваться новая рамка, а в ее поле выбирается новое изображение лица. Таким образом, эталон 1 используется только один раз — при первой оценке величины  $R$ . Результаты первого вычисления расстояния  $R$  для трех разных изображений с камеры представлены на рис. 4.2.2.

Второй и третий случай здесь отличаются от первого сильным приближением и соответственно удалением лица от объектива камеры.

Нормальное положение лица у объектива камеры соответствует первому случаю, при котором  $R = S = 29$ .

Заметим, что хотя во втором случае первый выбор изображения лица был не совсем удачным (образ не отцентрирован в поле рамки), тем не менее центры глаз определены достаточно точно. При этом расстояние между центрами глаз составило  $R = 33$  пикселя.

Размеры всех последующих новых изображений лиц будут определяться размерами новой рамки. Для приведения полученных при этом изображений к размеру  $112 \times 92$  пикселя далее используется процедура перемасштабирования. Таким образом, изменяя размер рамки и выбирая с ее помощью изображение лица (большее или меньшее) из исходного изображения и перемасштабировывая его, можно согласовать требуемые размеры.



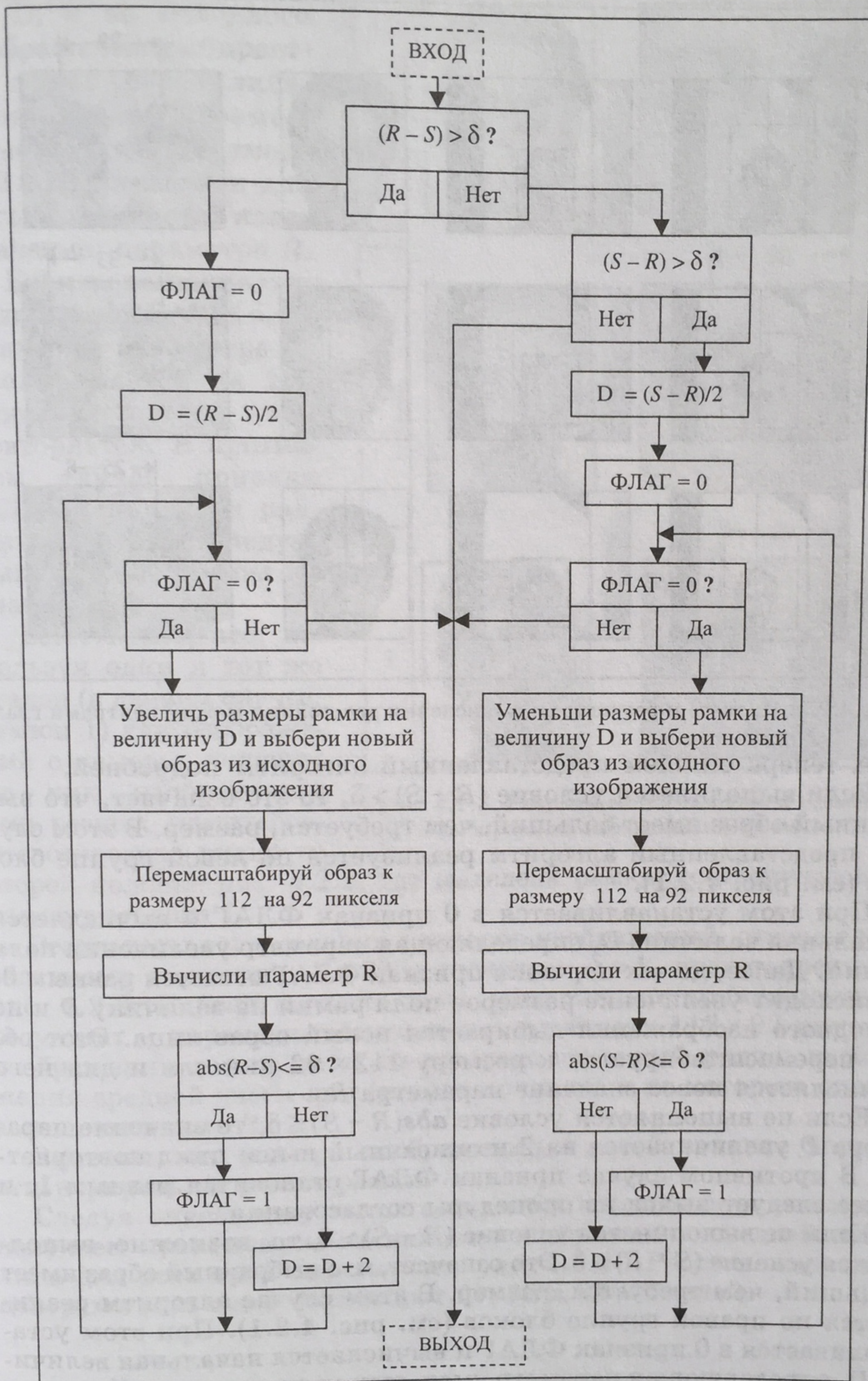


Рис. 4.2.1. Алгоритм согласования размера выделяемого образа



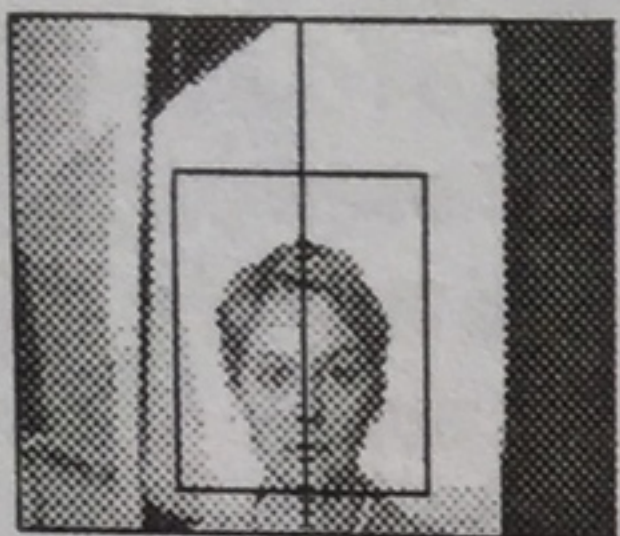


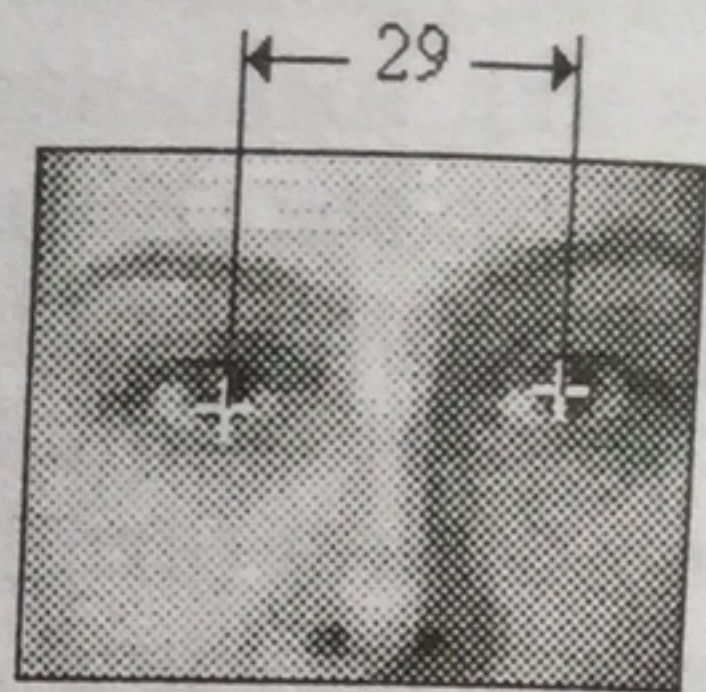
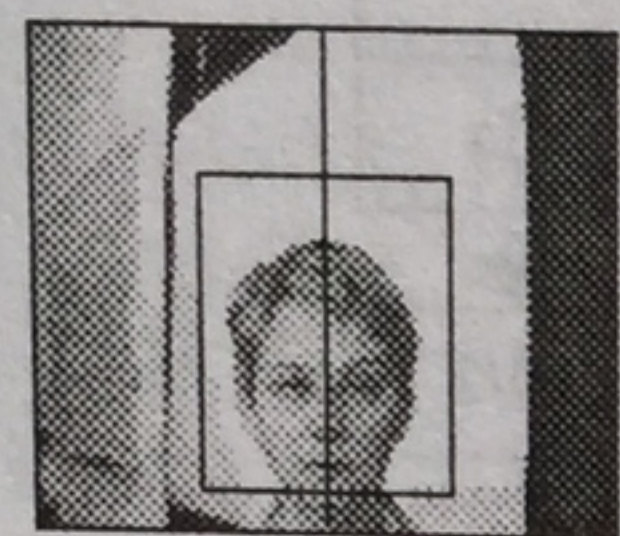


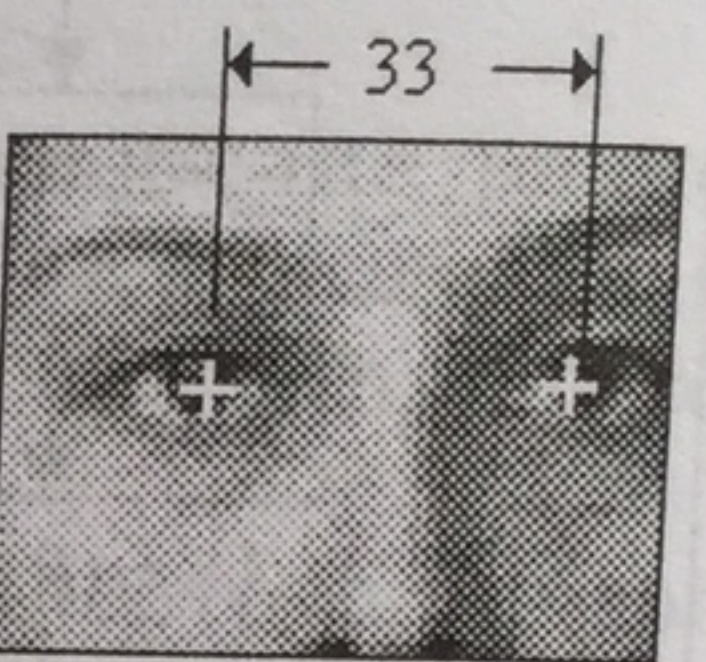
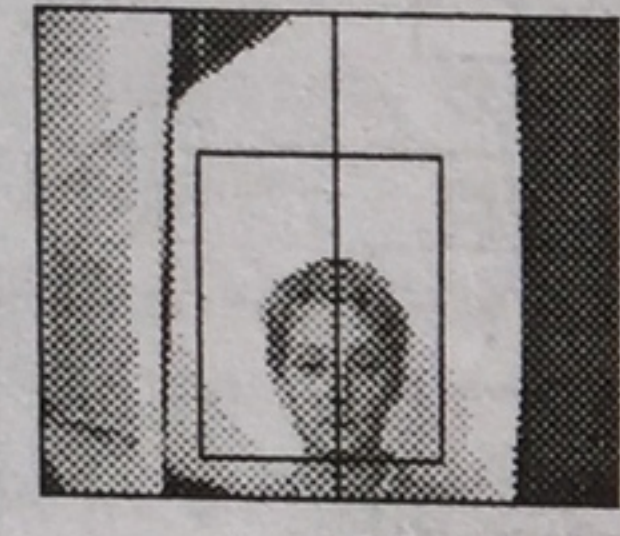
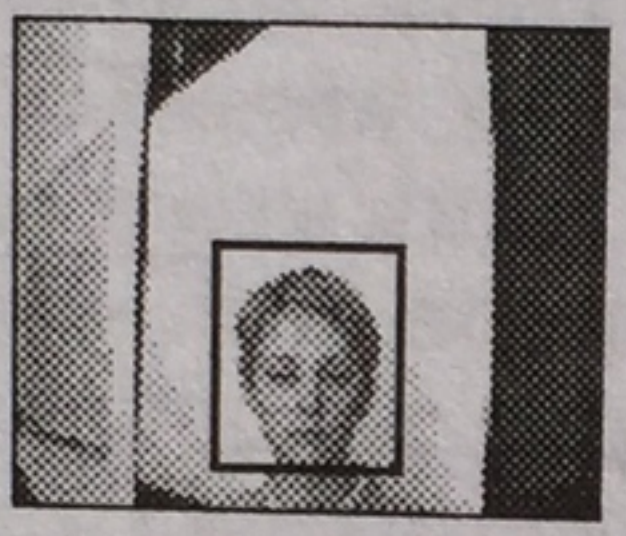
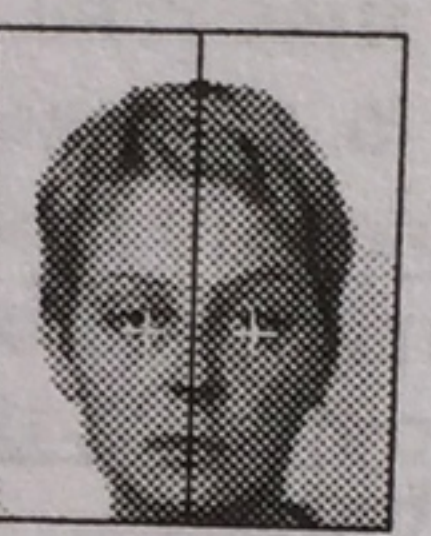
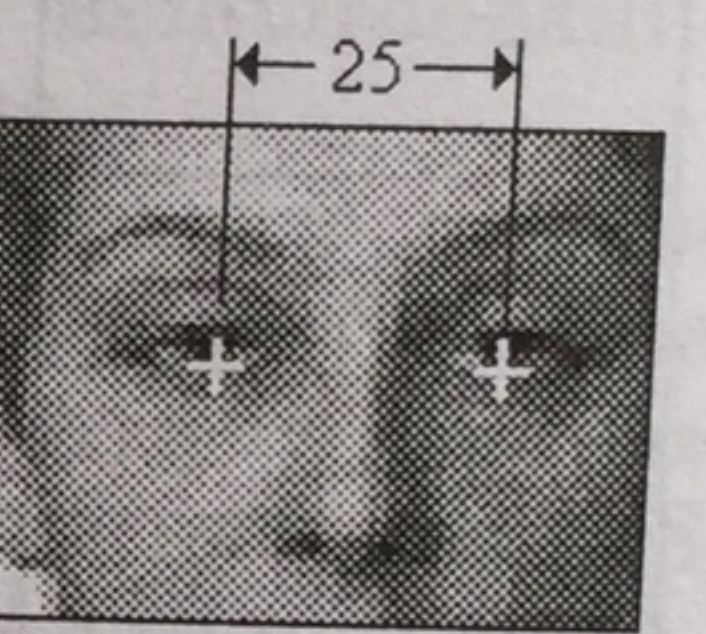
|   | Образ с камеры  | Первый выбор   | Выбранный образ   | Оценка расстояния $R$   |
|---|---|--|---|---|
| 1 |    |    |    |    |
| 2 |   |   |   |   |
| 3 |  |  |  |  |

Рис. 4.2.2. Процесс и результаты вычисления расстояния между центрами глаз

А теперь опишем представленный алгоритм подробнее.

Если выполняется условие  $(R - S) > \delta$ , то это означает, что выбранный образ имеет больший, чем требуется, размер. В этом случае представленный алгоритм реализуется по левой группе блоков (см. рис. 4.2.1).

При этом устанавливается в 0 признак ФЛАГ и вычисляется начальная величина  $D$ , определяющая параметр увеличения поля рамки. Далее, до тех пор пока признак ФЛАГ остается равным 0, происходит увеличение размеров поля рамки на величину  $D$  и из исходного изображения выбирается новый образ лица. Этот образ перемасштабируется к размеру  $112 \times 92$  пикселя и для него вычисляется новое значение параметра  $R$ .

Если не выполняется условие  $abs(R - S) \leq \delta$ , то значение параметра  $D$  увеличивается на 2 и описанный выше цикл повторяется. В противном случае признак ФЛАГ становится равным 1, и далее следует выход из процедуры согласования.

Если не выполняется условие  $(R - S) > \delta$ , то, возможно, выполняется условие  $(S - R) > \delta$ . Это означает, что выбранный образ имеет меньший, чем требуется, размер. В этом случае алгоритм реализуется по правой группе блоков (см. рис. 4.2.1). При этом устанавливается в 0 признак ФЛАГ и вычисляется начальная величина  $D$ , определяющая параметр уменьшения поля рамки. Далее, до тех пор пока признак ФЛАГ остается равным 0, происходит



уменьшение размеров поля рамки на величину  $D$ , и из исходного изображения выбирается новый образ лица. Этот образ перемасштабируется к размеру  $112 \times 92$  пикселя и для него вычисляется новое значение параметра  $R$ .

Если не выполняется условие  $abs(S-R) \leq \delta$ , то значение параметра  $D$  увеличивается на 2 и описанный выше цикл повторяется. В противном случае признак ФЛАГ становится равным 1 и далее следует выход из процедуры согласования.

Естественно, что, используя один и тот же эталон (в данном случае, эталон 1) для изображений с камеры, получим из них изображения лиц разных масштабов. Это отчетливо видно во второй колонке рис. 4.2.2, где выделена рамка, соответствующая размеру эталона 1.

Результаты согласования размеров изображений, отвечающих заданному параметру  $S = 29$  и значению  $\delta = 1$ , представлены на рис. 4.2.3. Первая колонка здесь представляет промежуточный результат — перемасштабированные к размеру  $112 \times 92$  пикселя изображения целого лица, из которых затем выбирались изображения средней части лиц (вторая колонка) и по которым оценивалась величина  $R$ .

Заметим, что все три изображения в левой колонке хорошо отцентрированы в поле рамки.

Следуя описанному выше алгоритму согласования размеров, получаем практически одинаковые изображения средней части лица для всех трех случаев (см. рис. 4.2.3). Расстояния между центрами глаз здесь составляют 29, 30 и 29 пикселей.

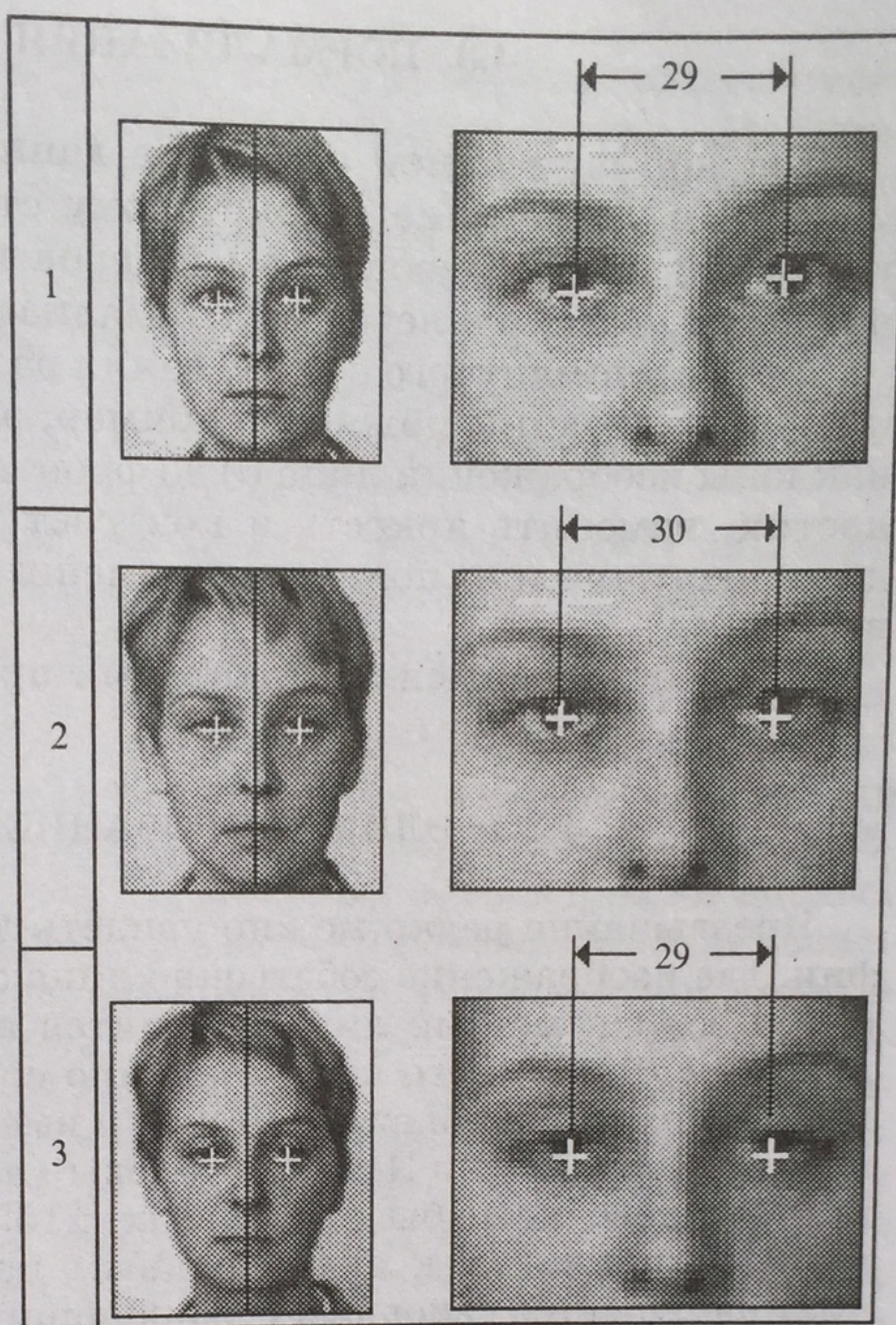


Рис. 4.2.3. Результаты согласования размеров выбираемых изображений с камеры



### 4.3. НОРМАЛИЗАЦИЯ ОБРАЗОВ

Под нормализацией образов в данном случае будем подразумевать их приведение к некоторому стандартному виду. В некотором смысле согласование размеров изображений, рассмотренное выше, тоже является их нормализацией.

Кроме упомянутого согласования размеров изображений в рамках нормализации можно, например, элиминировать повороты и наклоны изображений лица (!) во фронтальной и продольных плоскостях, изменить яркость и контраст исходного изображения и даже частично или полностью изменить его способ и условия освещения.

Ниже остановимся на некоторых процедурах нормализации.

#### 4.3.1. ЭЛИМИНИРОВАНИЕ ПОВОРОТОВ

Чрезвычайно редко можно увидеть лицо человека на фотографии, где изображение собственно лица ориентировано строго вперед, а ось симметрии лица находится в строго вертикальном положении. Чаще всего при этом лицо слегка повернуто, как и голова, вверх/вниз или вправо/влево и имеет тот же наклон в сторону плеч, что и голова. Поворот головы (например, вверх/вниз или вправо/влево хотя бы в пределах  $\pm 15^\circ$ ) в момент съемки лица камерой приводит к значительным изменениям его изображения: нарушению основных его пропорций, и, как следствие этого, изменению характеристик признаков и соотношений антропометрических параметров. Это становится особенно заметным на центральной части изображения лица (см., например, рис. 4.1.4b), что не позволяет принимать ее в данном случае за наиболее информативную часть распознаваемого образа.

С теоретической точки зрения, проблема элиминирования (учета, компенсации) поворота головы на изображениях лиц решена. При этом используются два основных подхода, представленные в работах [Adi97, Alf99, Ben98, Che94, Kar98, Kuc99, Lan97, Lam98, Smi99, Tha98, Vet97]: путем выделения признаков, являющихся инвариантами поворотов, либо путем составления виртуальных моделей лиц.

Первый подход [Alf99, Kuc99] реализуется путем вычисления контуров лиц, их проецирования на плоскость XY, нормирования и выравнивания масштабов контуров по осям X и Y, а также использования инвариантов типа Affine Arc Length, Enclosed Area, UNL и др. При этом далее в процессе распознавания используются только эти инварианты. Второй подход — синтез виртуальных моделей лиц — может быть реализован несколькими способами. Об одном из них — эластичных моделях лиц [Lan97, Smi99] — мы уже упоминали в разд. 2.3.6. Рассмотрим еще два новых подхода.

В первом из них [Ben98, Che94] определяются углы — параметры (азимут — угол поворота головы влево/вправо и элева-



ция — угол поворота головы вверх/вниз) поворота головы путем корреляционного анализа срезов для спектров Фурье—Меллина исходного и эталонного изображения лица, реконструкции нового положения головы на основе базовых моделей, вычисленных углов и синтеза нового изображения лица. Синтез осуществляется либо на основе обратного преобразования Фурье—Меллина модифицированных в соответствии с вычисленными углами спектров, либо на основе подбора базовых моделей лиц, отвечающих вычисленным углам—параметрам поворотов.

Синтез виртуальных моделей [Lam98, Tha98, Vet97] лиц основан на сравнении значений признаков заданного изображения со значениями признаков эталонного изображения лица (например, для изображения строго в фас), согласовании (замене) необходимых параметров и реконструкции нового изображения лица с учетом измененных параметров. Для успешного решения этой задачи необходимо прежде всего иметь не менее девяти эталонных изображений лиц в каждом классе: фас, поворот влево, поворот вправо, поворот вверх, поворот вниз и еще, по крайней мере, четыре совместных поворота — вниз и влево, вниз и вправо, вверх и влево и вверх и вправо. Далее необходимо так выбрать пространство исходных признаков и так его редуцировать, чтобы хотя бы два признака редуцированного пространства признаков «отвечали» за поворот лица: один за поворот влево/вправо, другой за поворот вверх/вниз. При этом другие признаки должны «отвечать» за индивидуальные особенности каждого класса образов. После этого можно привести интересующие нас два признака анализируемого образа к средним значениям, не меняя остальные, что будет соответствовать изображению лица, полученному строго в фас, но сохранившему индивидуальные особенности класса. Далее обычным путем завершается решение задачи классификации, а если требуется, то на основе обратных преобразований (например, Карунена—Лозва) по новым признакам реконструируется новое изображение лица. При этом, если изменять значения признаков, отвечающих за индивидуальные особенности класса, то можно от образа лица одного класса перейти к образу лица другого класса. А поскольку такого изображения (такого конкретно образа) могло в базе не быть, подобный образ и называется *виртуальным*.

Заметим, что ввиду сложности представленных процессов элиминирования поворотов головы на изображениях лиц эта проблема не нашла широкого практического применения в системах распознавания «on-line».

Однако, если система «on-line» все же ориентирована на распознавание лиц с учетом поворотов головы, то реализуется это за счет наполнения базы данных, которая должна содержать в каждом классе необходимое множество образов, соответствующих различным поворотам. Примером может служить база данных [Orl\*\*], хотя и здесь упомянутое выше требование выполнено не для всех классов образов.



### 4.3.2. ЭЛИМИНИРОВАНИЕ НАКЛОНОВ

В системах распознавания «on-line» элиминирование наклонов лица на изображениях не всегда необходимо. Малые углы наклона система распознавания вообще не должна как бы замечать — быть по отношению к ним робастной и, следовательно, не должна затруднять процесс распознавания и ухудшать его качество. При больших углах наклона, например больших чем  $30^\circ$ , необходимо просто сообщить пользователю о непригодности исходного изображения и повторить съем нового изображения с камеры.

Проанализируем теперь примеры, для которых возможно точное определение угла наклона головы простыми методами. Далее опишем эти методы и приведем конкретные результаты элиминирования наклонов в системе «on-line».

На рис. 4.3.1 приведены два типичных варианта наклонов лиц, полученные с исходных изображений непосредственно с камеры. Ясно, что они не пригодны для распознавания, хотя наибольший угол наклона здесь  $20^\circ$ .

Более того, если оценивать расстояния между центрами глаз по выделенным частям лиц, то не исключены существенные ошибки в оценках: во-первых, эти расстояния будут меньше, чем реальные, а во-вторых, для их корректировки необходимо будет учесть, в какую из сторон был выполнен наклон головы — влево или вправо. Неверное же определение оси симметрии лица не позволит точно центрировать выделяемый образ в поле рамки, что также осложнит весь процесс распознавания и повлияет на его результат.

Обозначим центр левого (по отношению к нам) глаза цифрой 1, а центр правого (также по отношению к нам) глаза цифрой 2 и обозначим параметры, которые необходимо знать, чтобы вычислить угол наклона головы.

Для этого выполним построение взаимного расположения глаз для двух типичных вариантов наклона (рис. 4.3.2), где  $R$  — расстояние между центрами глаз,  $D$  — расстояние между высотами центров глаз,  $\theta$  — оцениваемый угол.

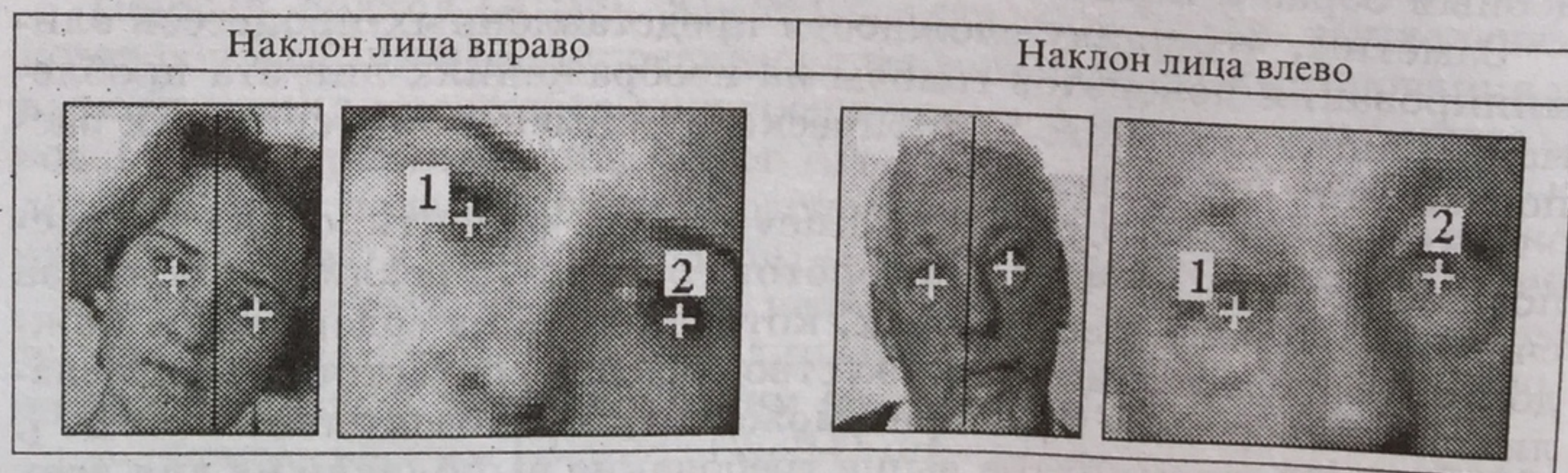


Рис. 4.3.1. Два типичных варианта наклонов лиц на исходных изображениях



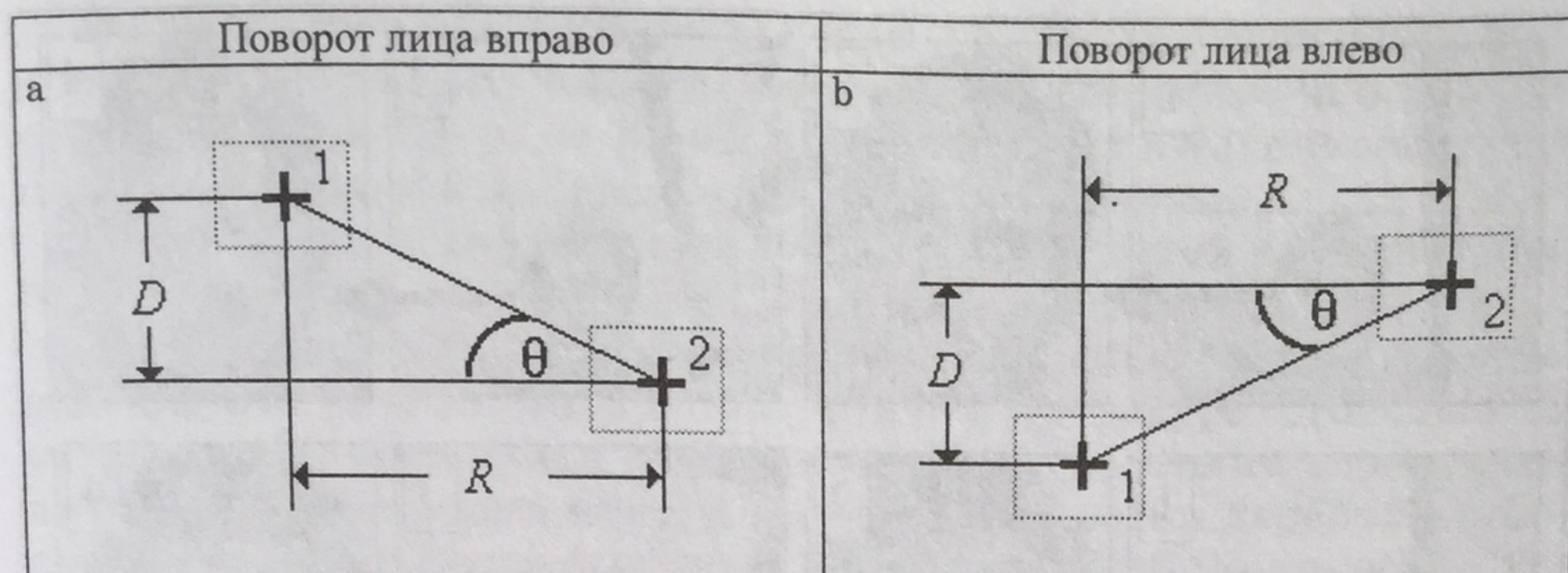


Рис. 4.3.2. Определение величин  $R$ ,  $D$  и  $\theta$  для поворотов лица во фронтальной плоскости: а — поворот лица вправо; б — поворот лица влево

Легко заметить, что угол

$$\theta = \arctg \frac{D}{R} \quad (4.1)$$

и может принимать значения как отрицательные (что соответствует повороту головы вправо), так и положительные (что соответствует повороту головы влево).

Оценим теперь абсолютную величину угла  $\theta$  в зависимости от значений величины  $D$ , приняв, например, что расстояние между центрами глаз  $R = 30$  пикселей. Такое расстояние соответствует изображению лица размером  $112 \times 92$  пикселя, в котором собственно лицо размещено в поле всего изображения.

Зависимость угла  $\theta$  от величины  $D$  при фиксированном значении  $R = 30$  показана на рис. 4.3.3. Из него видно, что для исходных изображений размером  $112 \times 92$  пикселя угол  $\theta = 20^\circ$  практически получается уже при разнице в высотах центров глаз  $D = 10$  пикселей. Угол  $\theta < 5^\circ$ , который можно не учитывать и, соответственно, не элиминировать, соответствует значению  $D$ , равному 2 пикселя.

Оцененные граничные значения  $D$  могут служить в дальнейшем критерием для принятия решения об элиминировании наклонов головы в реальных ситуациях.

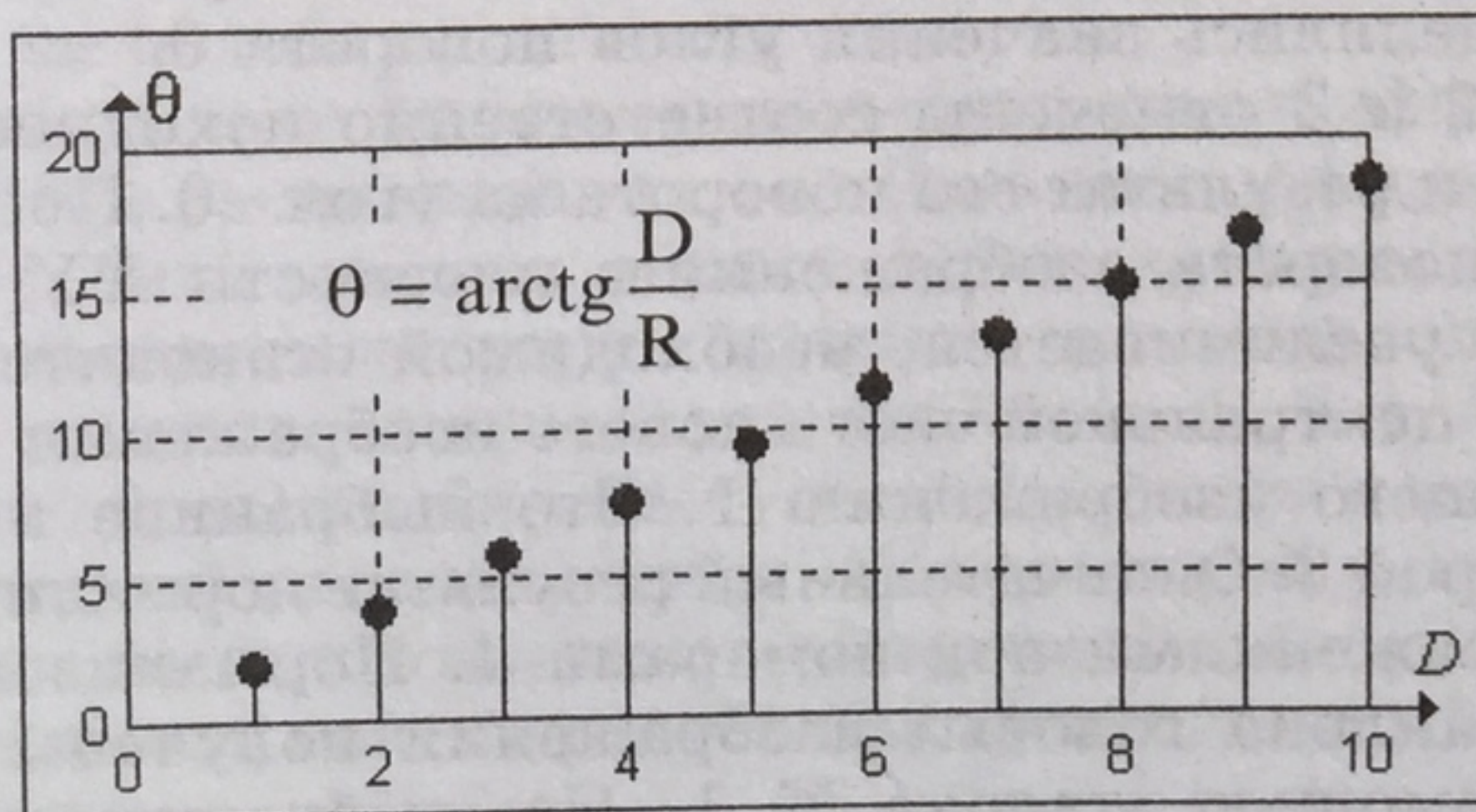


Рис. 4.3.3. Зависимость угла  $\theta$  от величины  $D$  при фиксированном значении  $R = 30$



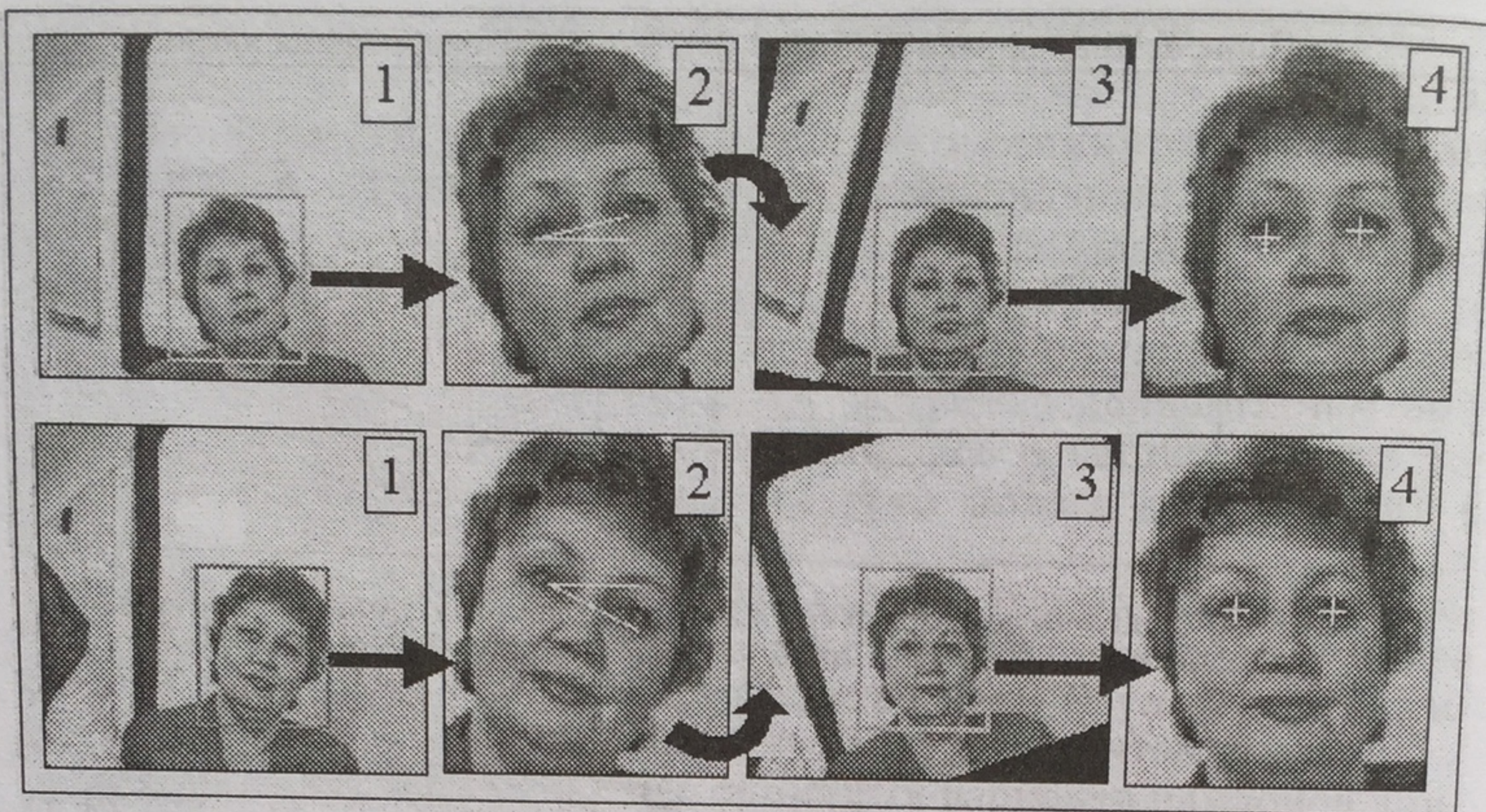


Рис. 4.3.4. Процесс вычисления углов  $\theta$  и нормализации изображений

Весь процесс вычисления углов  $\theta$  и нормализации изображений лиц показан на рис. 4.3.4. Заметим, что для правильной нормализации (поворота изображений в плоскости  $XU$ ) используется угол со знаком, противоположным вычисленному.

В первом случае (см. рис. 4.3.4) угол  $\theta$  составил  $11^\circ$ , а поворот исходного изображения с камеры был выполнен для  $\theta = -11^\circ$ . Во втором случае угол  $\theta$  составил соответственно  $-23$  и  $23^\circ$ .

На рис. 4.3.4 цифрами 1 отмечены исходные изображения, полученные с камеры. На них рамками обозначены регионы, в которых были обнаружены методом сопоставления с эталоном, описанным выше, образы лиц. В обоих случаях использовался один и тот же эталон № 1 (см. рис. 4.1.2). Цифрами 2 отмечены изображения выделенных лиц, из которых затем с помощью эталона № 3 (см. рис. 4.1.2) были получены изображения средней части лиц. По ним с использованием эталонов глаз определялось положение левого и правого глаза, вычислялись центры глаз и далее определялся угол наклона линии глаз относительно горизонтали. Таким образом определялись значения углов поворота  $\theta$ .

Цифрами 2 и 3 отмечены соответственно исходное изображение с камеры и результат его поворота на угол  $-\theta$ . Поскольку при выполнении поворота изображения в плоскости  $XU$  его размер, как правило, увеличивается, необходимой становится операция выбора из его центральной части нового изображения по размеру, соответствующего изображению 1. Это выбранное изображение отмечено цифрой 3. Окончательный результат нормализации представлен изображениями под номерами 4. Нормализованные (относительно наклона головы) изображения получены из изображения 3 с помощью эталона № 1. На изображении 3 в рамке выделено поле, в котором среднеквадратическое отклонение от эталона было наименьшим.



Из представленных результатов видно, что нормализация прошла успешно: выделенные изображения лиц достаточно хорошо отцентрированы в поле рамок (размером  $112 \times 92$  пикселя) и в них элиминирован первоначальный наклон головы.

Дальнейшая обработка полученных изображений в системе «on-line» не должна вызывать практических трудностей.

Используя теперь в качестве исходных изображения, отмеченные цифрой 4а и 4б, можно приступить к процедуре согласования их размеров, вычислить расстояние между центрами глаз, выделить все необходимые классы подобразов, а затем перейти к процедуре нормализации выделенных изображений по яркости. Эта последняя процедура нормализации изображений будет описана в следующем разделе.

#### 4.3.3. НОРМАЛИЗАЦИЯ ЯРКОСТИ ИЗОБРАЖЕНИЙ

В некоторых системах распознавания «on-line» в качестве камер, предназначенных для получения исходных изображений, используются видеокамеры не очень высокого качества (например, предназначенные для видеоконференций). Главная проблема, возникающая при этом, — нестабильность яркостных параметров изображения: яркость и контраст изображения полностью зависят от окружающего (наружного) освещения.

Другая проблема возникает в том случае, когда исходные изображения лиц людей из разных классов собирались в базу данных из различных источников и имеют, естественно, разные уровни яркости, контраста и прорисовки отдельных деталей на участках в области лиц и окружающего их фона.

Нестабильность или различие в уровнях яркости становятся особенно опасными для методов распознавания, в которых исходными признаками являются значения пикселей. И хотя в таких методах с помощью яркости выявляются форма и детали образа (например, овал лица, форма рта, бровей, глаз), система будет скорее реагировать на подобие в освещении некоторого эталона и исходного изображения, чем на подобие в деталях лиц.

Чтобы избавиться от такого нежелательного эффекта, дистрибьюторы систем распознавания «on-line» рекомендуют, например, записывать в базы данных образы при различном наружном освещении: естественном днем и вечером, с использованием дополнительного искусственного освещения и т. д., а также локально (под каждый образ) подбирать чувствительность камеры в момент сеанса с системой распознавания. Смысл такой рекомендации — иметь в базе данных эталоны «на все возможные случаи» (!) и, возможно, нивелировать яркость как существенный признак при распознавании.

Одним из методов улучшения таких изображений является видоизменение гистограммы [Прэ82] — преобразование яркостей



исходного изображения так, чтобы гистограмма распределения яркостей выходного изображения приняла желаемую форму. Применительно к задачам распознавания лиц конкретные рекомендации для реализации этого и других методов можно найти в работах [Adi97, Alf99, Lan97, Tsa98].

На рис. 4.3.5 с левой стороны представлены два изображения, снятые с камеры в различных условиях освещения. На верхнем, более темном, изображении в поле рамки едва различимы детали лица. На нижнем изображении, полученном при более или менее нормальном освещении, детали лица в поле рамки различаются значительно лучше.

Однако как оценить, достаточно ли качество этого нижнего изображения для задач распознавания и не нужно ли (и можно ли) его еще улучшить. Обратим внимание на гистограммы, соответствующие изображениям, находящимся в поле рамок. Как видно, гистограмма распределения яркостей верхнего, более темного, изображения «сдвинута» влево (не использован диапазон уровней начиная со 150 и до 255), в то время как гистограмма нижнего лучшего качества: она занимает практически весь диапазон яркостей. Однако видно, что и здесь не полностью использован диапазон яркостей на уровнях от 0 до 25.

При недоиспользовании диапазона высоких уровней яркости получается слишком темное изображение, в котором трудно вообще выделить детали лиц, в то время как при неиспользовании диапазона низких уровней яркости границы черного и белого то-

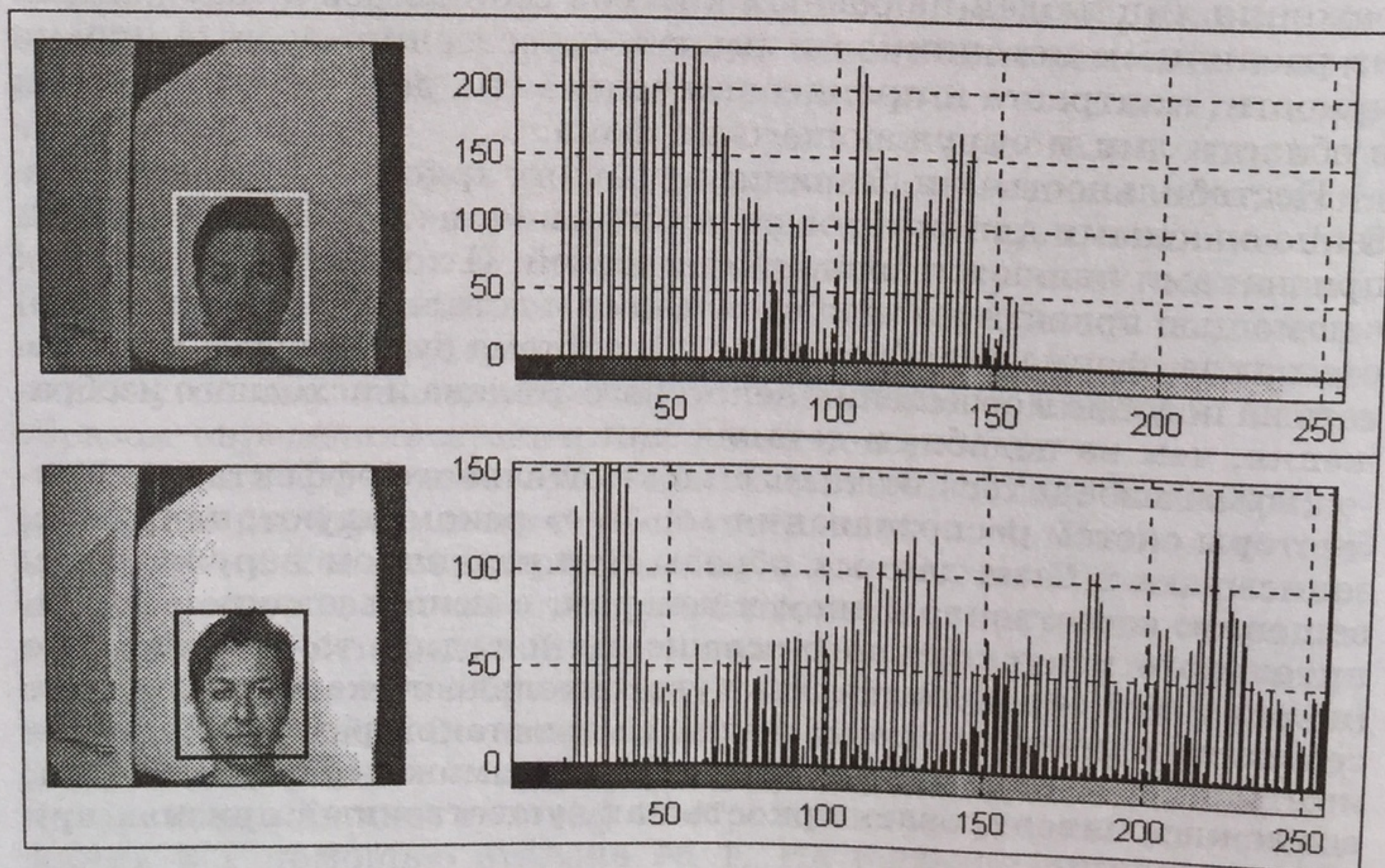


Рис. 4.3.5. Изображения с камеры и гистограммы распределения яркости выделенных образов



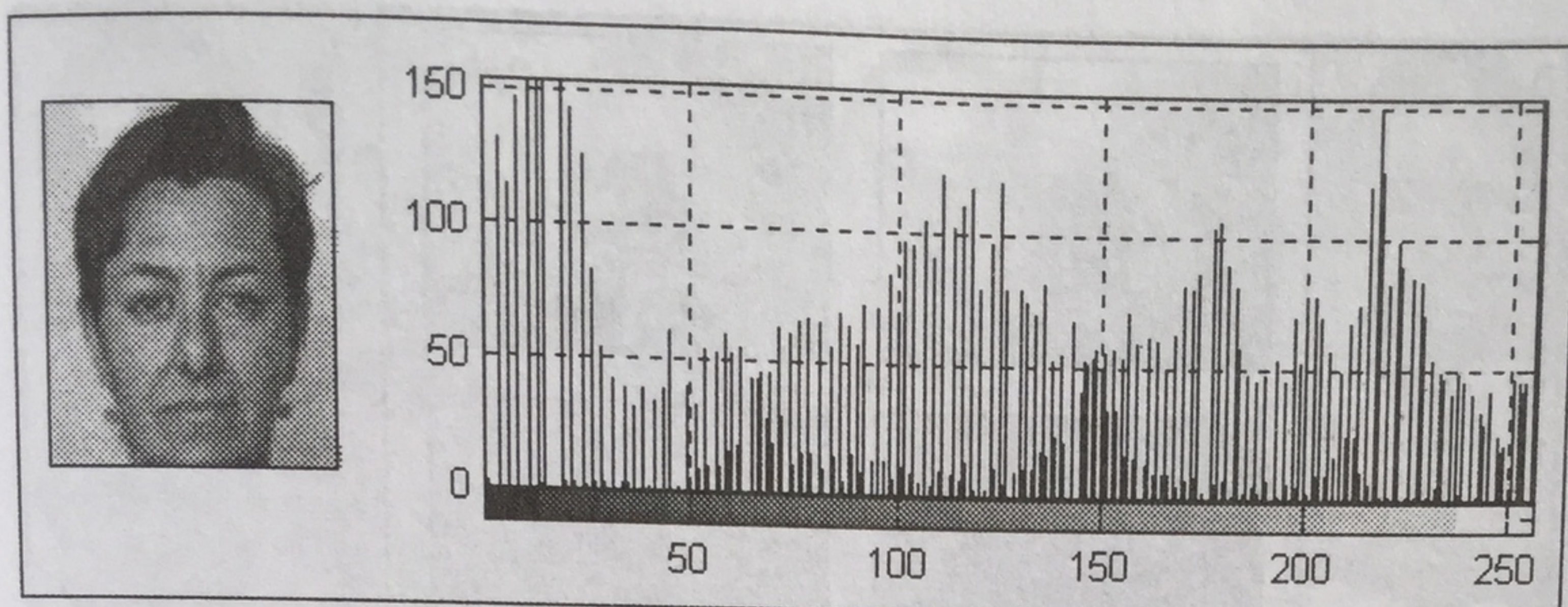


Рис. 4.3.6. Изображение после обработки и его гистограмма распределения яркостей

нов становятся нерезкими, отдельные детали лиц остаются непро-рисованными и т. д.

Изменим яркость нижнего изображения (в рамке) так, чтобы гистограмма распределения яркостей занимала весь диапазон от 0 до 255. Результат такого изменения изображения и соответствующая ему гистограмма представлены на рис. 4.3.6. Хотя «на глаз» не видно существенных изменений в целом на изображении, но при детальном рассмотрении лица можно заметить, что изменения есть.

В автоматическом режиме подобная операция нормирования яркости изображения реализуется следующим образом:

$$G_{ij} = 255(Z_{ij} - Z_{\min}) / (Z_{\max} - Z_{\min}), \quad (4.2)$$

где  $Z_{ij}$  — пиксель исходного изображения,  $i = 1, 2, \dots, M$  и  $j = 1, 2, \dots, N$ ;  $M, N$  — размеры изображения;  $Z_{\min}$  — минимальное значение яркости исходного изображения;  $Z_{\max}$  — максимальное значение яркости исходного изображения;  $G_{ij}$  — пиксель изображения с видоизмененной гистограммой.

В некоторых случаях, связанных с обработкой исходных изображений, полученных из разнородных источников, перед нормированием (4.2) исходное изображение центрируется относительно своего среднего значения и нормируется на среднеквадратическое отклонение от среднего.

В системах распознавания «on-line» операцию (4.2) нормирования необходимо применять дважды: первый раз непосредственно к исходному изображению, полученному с камеры, что создаст лучшие условия работы блоков выделения изображений по эталонам; второй раз — к выделенному изображению, поскольку оно является локальным по отношению ко всему исходному изображению и в нем не использован весь диапазон изменения яркости.

Результат нормализации (4.2) исходных изображений, полученных с камеры в различных условиях наружного освещения, представлен на рис. 4.3.7.



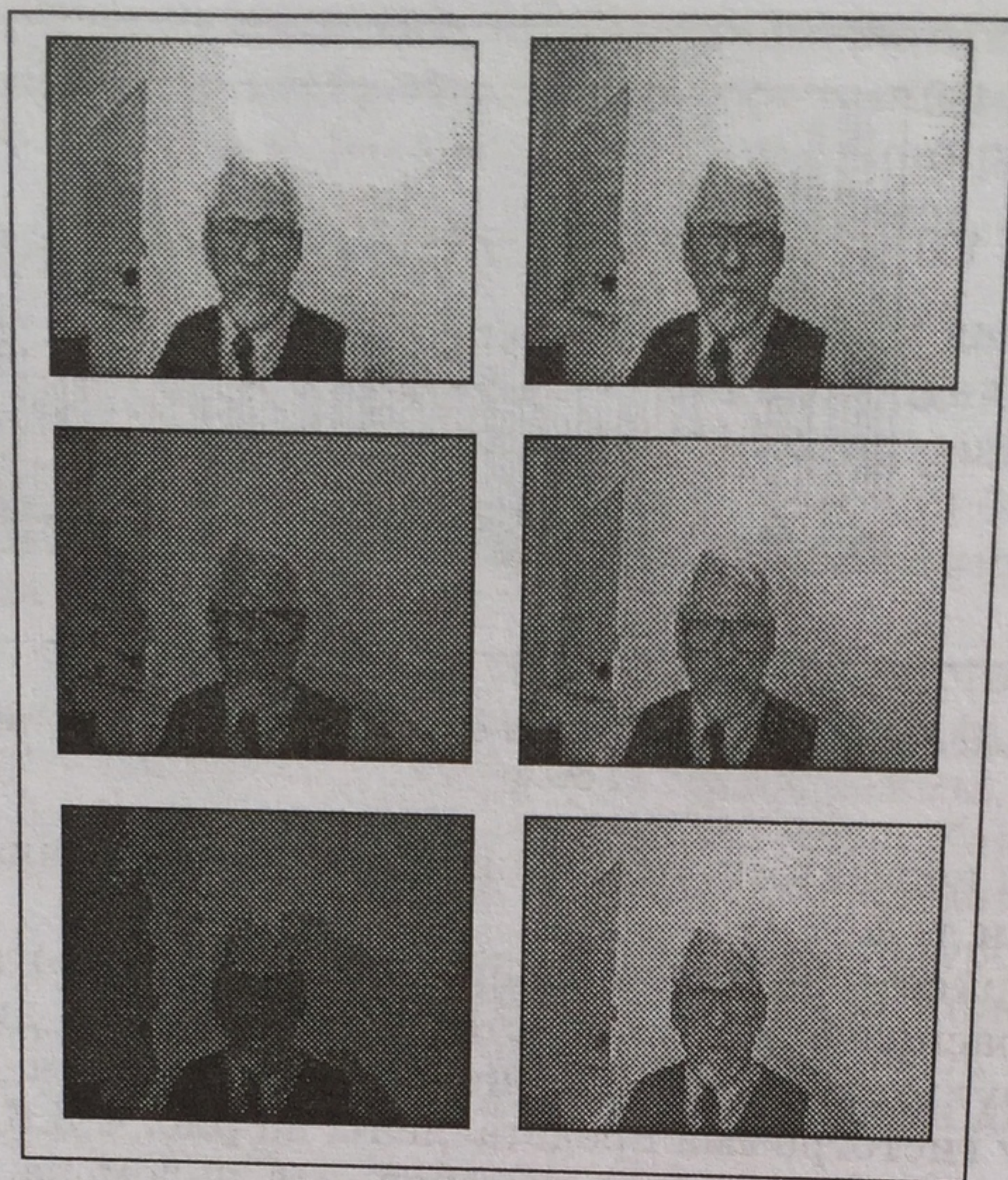


Рис. 4.3.7. Нормализация яркости изображений, полученных с камеры

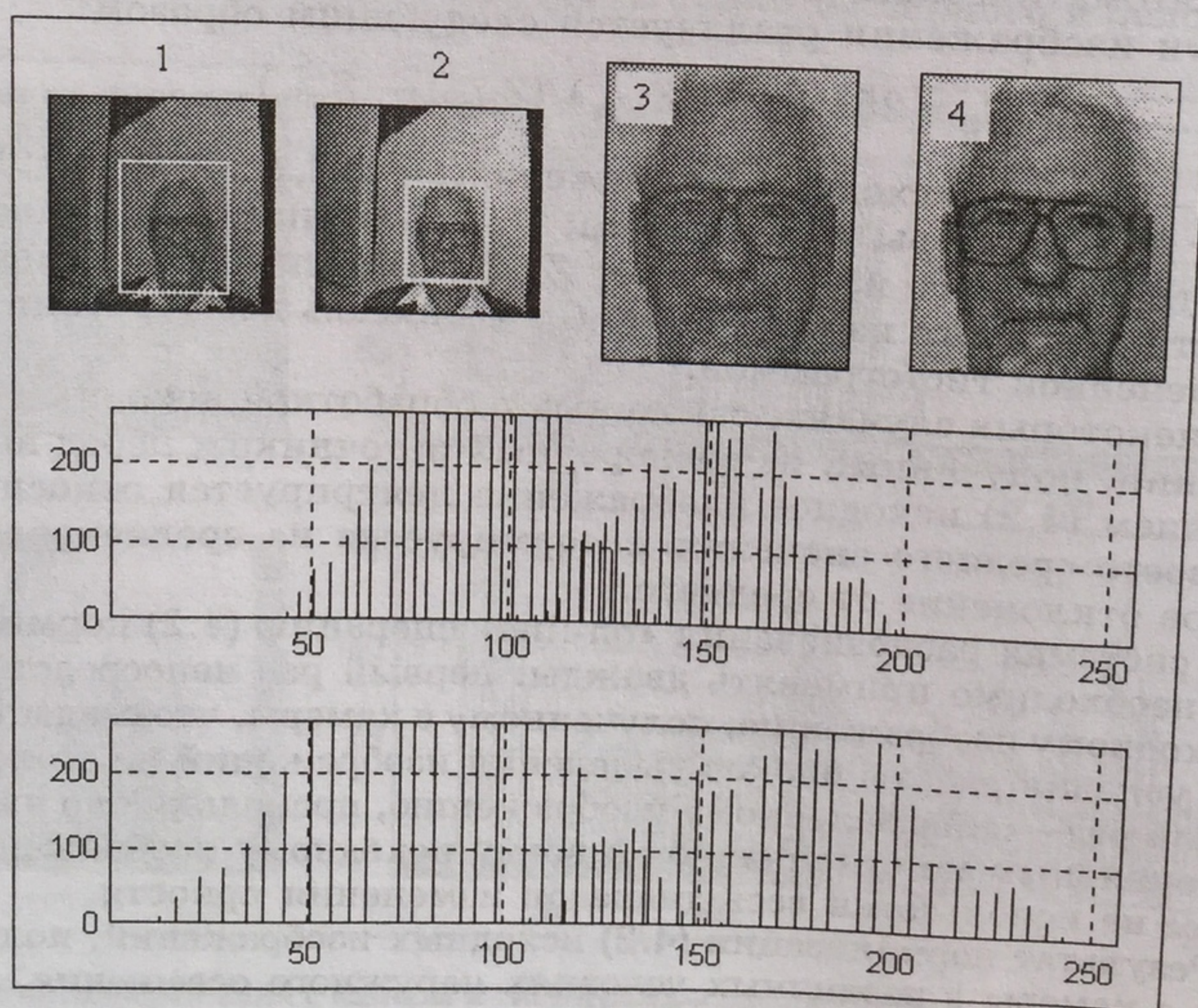


Рис. 4.3.8. Результаты нормализации яркости по двухэтапной процедуре





Рис. 4.3.9а. Изображения, полученные из разнородных источников и при разных условиях освещения в момент съемки



Рис. 4.3.9б. Результаты нормализации яркости для изображений базы данных, полученной из разнородных источников



Если для верхнего изображения нормализация яркости, возможно, и не была необходимой, то для двух нижних изображений она дала значительный эффект и, стало быть, была остро необходима.

А теперь проверим, насколько обязательным становится выполнение второй нормализации после выделения изображения лица. Выберем такой уровень наружного освещения, при котором уже после выполнения первой нормализации выделенное изображение лица имело бы достаточную яркость. Далее построим гистограмму распределения яркости этого изображения, нормализуем его и построим новую гистограмму. Сравнивая две гистограммы, оценим необходимость выполнения второй нормализации.

Результаты эксперимента, выполненного по двухэтапной процедуре нормализации яркости, представлены на рис. 4.3.8.

Здесь изображения 1 и 2 — исходные, полученные с камеры и нормализованные по яркости. Изображение 3 — результат выделения из поля рамки по изображению 2. Изображение 4 получено из изображения 3 путем нормализации его по яркости. Верхняя гистограмма относится к изображению 3, а нижняя соответственно к изображению 4. Как видим, этот пример, в частности, убедительно показывает необходимость двухэтапной процедуры нормализации изображений для систем распознавания «on-line».

Целью дальнейшей нормализации яркости рассматриваемого класса изображений должно стать подчеркивание границ форм лица: его овала, границ рта, глаз, бровей и т. д. Для этого можно воспользоваться известными методами [Пре82], основанными, например, на локальных методах обработки изображений.

В заключение настоящего раздела представим результаты нормализации яркости изображений (из базы данных), полученных из разнородных источников. На рис. 4.3.9а представлены исходные изображения, на рис. 4.3.9b — изображения, полученные в результате нормализации. Сравнение результатов говорит само за себя и в пользу нормализации.

Еще раз подчеркнем: качество распознавания образов в огромной степени зависит от качества самих образов. Именно поэтому улучшению качества образов, предназначенных для распознавания, необходимо уделить особое внимание. Все процедуры предварительной обработки изображений, рассмотренные в настоящем разделе, призваны способствовать улучшению их качества, а также наилучшему выделению признаков вообще или биометрических признаков в частности. В связи с этим в рассматриваемые ниже системы распознавания включен (как обязательный) блок предварительной обработки изображений.



## 5. СИНТЕЗ СИСТЕМЫ РАСПОЗНАВАНИЯ ДЛЯ СТРУКТУРИРОВАННЫХ ДАННЫХ

Под структурированными данными будем понимать данные, сгруппированные в классы. В данном случае каждый класс данных — изображения лиц одного и того же человека. А поскольку эти данные сгруппированы, то каждое изображение лица является образом в классе. В общем случае число образов в каждом классе может быть разным, но для упрощения последующего изложения примем, что мы имеем дело с базами, в которых число образов в каждом классе одинаковое. Пример таких данных — база [Orl\*\*].

Под синтезом системы распознавания будем подразумевать некоторый процесс, в ходе которого выполняется:

- разделение исходных данных на обучающую и контрольную группы;
- выбор признаков исходного пространства признаков;
- редукция пространства признаков и анализ результата кластеризации в редуцированном пространстве признаков;
- реализация основной процедуры распознавания контрольных образов;
- анализ качества распознавания контрольных образов и их модификаций (наложение шума, масштабирование);
- анализ качества распознавания контрольных образов, не принадлежащих собственной базе данных;
- выбор параметров, обеспечивающих 100 %-ное распознавание контрольных образов.

Предлагаемая методика синтеза основана на подходах, представленных в разд. 2.4 и имеющих наивысшее качество распознавания (анализ соответствующих показателей качества смотри там же). Общий вид структурной схемы синтеза системы распознавания для структурированных данных представлен на рис. 5.1.1.

### 5.1. АВТОМАТИЧЕСКОЕ РАЗДЕЛЕНИЕ ДАННЫХ НА ОБУЧАЮЩУЮ И КОНТРОЛЬНУЮ ГРУППЫ

Итак, мы имеем исходную базу данных с числом классов  $K$  и числом  $Q$  образов в классе. Каждый образ имеет размер  $M \times N$  пикселей. Для синтеза системы распознавания необходимо выполнить следующие этапы (помечены знаком «точка»).

- В каждом классе вычисляем средний образ

$$\bar{O}_{M \times N}^{(k)} = \left( \sum_{q=1}^Q O_{M \times N}^{(k,q)} \right) / Q, \quad k = 1, 2, 3, \dots, K \quad (5.1.1)$$

и отклонение  $\Delta(q)$  от среднего  $\bar{O}_{M \times N}^{(k)}$  для каждого образа в классе



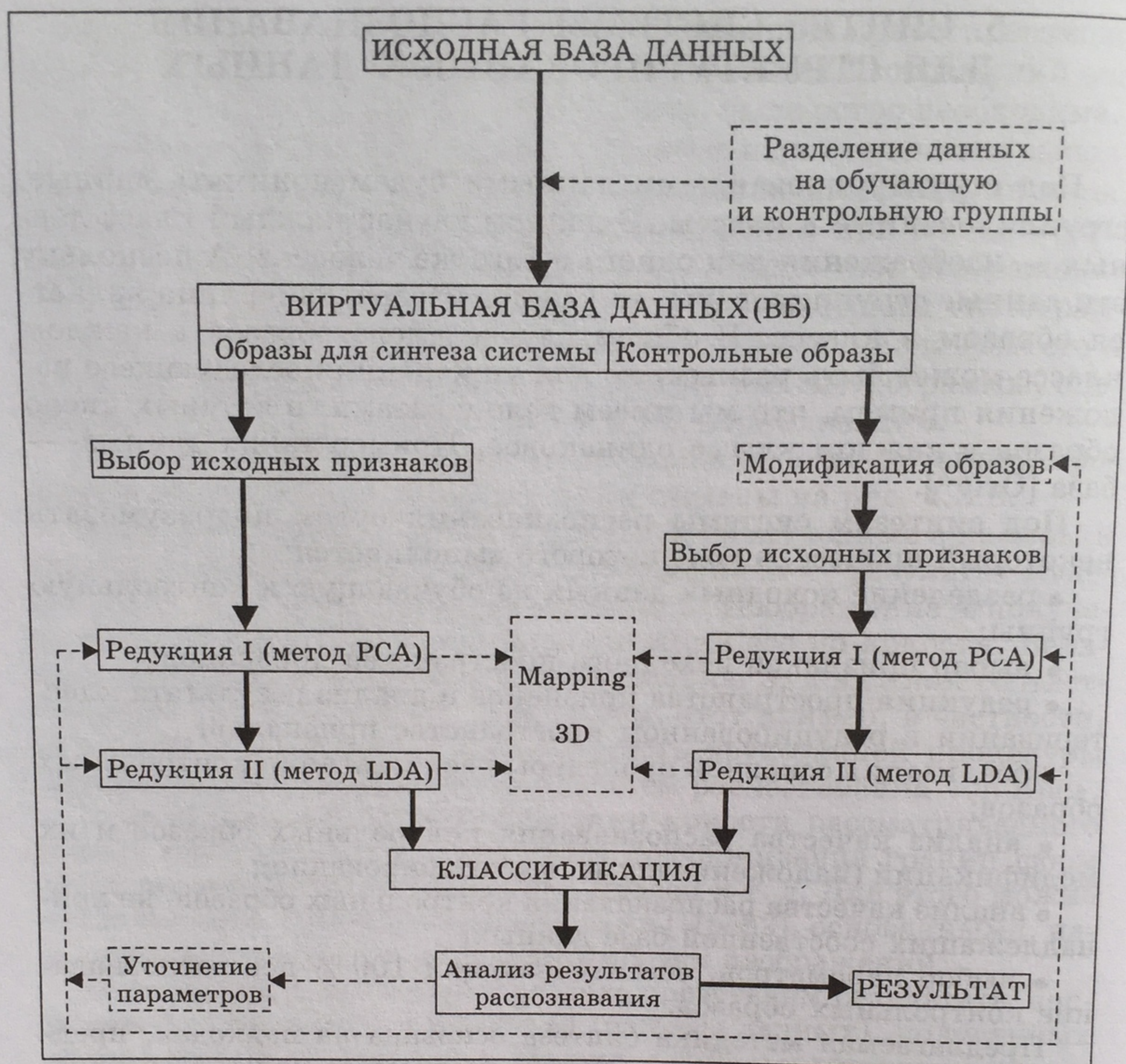


Рис. 5.1.1. Структурная схема синтеза системы распознавания

$$\Delta(q) = \sqrt{\left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N}^{(k)} \right)^2}, \quad q = 1, 2, 3, \dots, Q. \quad (5.1.2)$$

• Далее отсортируем значения  $\Delta(q)$  в порядке их убывания так, что  $\Delta(1) > \Delta(2) > \Delta(3) > \dots$ , и переупорядочим образы каждого класса исходной базы данных в соответствии с новыми порядковыми номерами отклонений. Назовем представленную таким образом базу образов «виртуальной базой».

• Выберем первые  $L$  образов каждого класса (причем  $L < Q$  и  $(Q - L) \geq 1$ ) в виртуальной базе (ВБ) для синтеза (построения/обучения) системы распознавания, а оставшиеся  $(Q - L)$  образов — для контроля распознавания.

Структура и параметры первой части ВБ представлены на рис. 5.1.2 слева, а второй части ВБ — справа. Здесь  $K$  — число классов;  $L$  — число образов для обучения;  $M, N$  — размеры образа.





Рис. 5.1.2. Структура виртуальной базы



Рис. 5.1.3. Порядок следования образов одного и того же класса в исходной (верхний ряд) и виртуальной базе данных (нижний ряд)

На рис. 5.1.3 в верхней части размещены образы одного из классов исходной базы данных [Orl\*\*], а в нижней — образы того же класса данных в виртуальной базе. Видно, что в ВБ последние четыре образа (под номерами 2, 9, 7 и 4) имеют наименьшие изменения, например в повороте головы, и, очевидно, наиболее близки к среднему образу в данном классе.

## 5.2. ОТБОР ПЕРВИЧНЫХ ПРИЗНАКОВ И ИХ РЕДУКЦИЯ

Исходные признаки образа — его яркость, т. е. значение каждого пикселя. Пространство всех признаков для ВБ имеет размер  $KL(MN)$  элементов. Цель настоящего раздела — представление



процедуры отбора первичных признаков и их редукции. В основе этих процедур лежат методы анализа главных компонент (РСА) и преобразования Карунена—Лоэва (ПКЛ, англ. KLT).

Для реализации поставленной задачи выполним следующие шаги [Pen00, Tsa98].

- На основе всех образов из первой части ВБ определим «средний образ»:

$$\bar{O}_{M \times N} = \frac{1}{LK} \sum_{k=1}^K \sum_{q=1}^L O_{M \times N}^{(k,q)}. \quad (5.2.1)$$

- Вычислим матрицы ковариаций  $R_{M \times M}$  и  $C_{N \times N}$ , из которых первая определяется относительно строк каждого централизованного образа, а вторая — относительно столбцов [Tsa98]:

$$R_{M \times M} = \frac{1}{LK} \sum_{k=1}^K \sum_{q=1}^L \left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N} \right) \cdot \left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N} \right)^T; \quad (5.2.2)$$

$$C_{N \times N} = \frac{1}{LK} \sum_{k=1}^K \sum_{q=1}^L \left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N} \right)^T \cdot \left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N} \right). \quad (5.2.3)$$

- Для каждой из матриц (5.2.2) и (5.2.3) определим собственные числа и соответствующие им собственные векторы, для которых выполняются условия

$$\left. \begin{aligned} \Lambda_{M \times M}^{(R)} &= \left[ V_{M \times M}^{(R)} \right]^T R_{M \times M} V_{M \times M}^{(R)}; \\ \Lambda_{N \times N}^{(C)} &= \left[ V_{N \times N}^{(C)} \right]^T C_{N \times N} V_{N \times N}^{(C)}, \end{aligned} \right\} \quad (5.2.4)$$

где  $\Lambda_{M \times M}^{(R)}$  и  $\Lambda_{N \times N}^{(C)}$  — диагональные матрицы, на диагоналях которых размещены соответствующие «строчному и столбцовому представлению» образов собственные числа;

$V_{M \times M}^{(R)}$  и  $V_{N \times N}^{(C)}$  — ортогональные матрицы, строки которых определяют собственные векторы, соответствующие собственным числам.

Диагонализация вида (5.2.4) возможна лишь в случаях, когда матрицы собственных векторов являются ортогональными, что можно записать в виде

$$\left. \begin{aligned} V_{M \times M}^{(R)} \cdot \left[ V_{M \times M}^{(R)} \right]^T &= I_{M \times M}; \\ V_{N \times N}^{(C)} \cdot \left[ V_{N \times N}^{(C)} \right]^T &= I_{N \times N}, \end{aligned} \right\} \quad (5.2.5)$$

где  $I_{M \times M}$ ,  $I_{N \times N}$  — единичные матрицы порядков  $M$  и  $N$ .

- В диагональных элементах матриц  $\Lambda_{M \times M}^{(R)}$  и  $\Lambda_{N \times N}^{(C)}$  найдем по



$p$  наибольших собственных чисел ( $p \ll MN$ ) и запомним порядок их следования.

- Из матрицы  $[V_{M \times M}^{(R)}]^T$  выберем  $p$  строк, а из матрицы  $V_{N \times N}^{(R)}$  —  $p$  столбцов, соответствующих  $p$  наибольшим собственным числам, и сформируем матрицы редукции признаков  $F_{p \times M}^{(R)}$  и  $F_{N \times p}^{(C)}$ .

- Для каждого образа из первой части ВБ выполним преобразование Карунена—Лоэва, которое может быть представлено в форме

$$Y_{p \times p}^{(k,q)} = F_{p \times M}^{(R)} \left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N} \right) F_{N \times p}^{(C)} \quad \begin{matrix} k = 1, 2, \dots, K; \\ q = 1, 2, \dots, L \end{matrix} \quad (5.2.6)$$

и реализовано на основе матричной схемы, представленной на рис. 5.2.1.

Пространство всех признаков имеет теперь размер  $KL(p^2)$  элементов, что меньше первоначального размера в  $MN/p^2$  раза. В матрице редуцированных признаков наибольшее значение имеют компоненты в верхнем левом углу.

Отображение редуцированного пространства признаков в системе координат 3D показано на рис. 5.2.2 (для параметров ВБ:  $K = 20$ ;  $L = 7$ ;  $p = 9$ ).

За координаты  $X$ ,  $Y$  и  $Z$  приняты наиболее значимые первые три компоненты для каждого образа, т. е. пиксели с адресами (1,1), (1,2) и (2,1).

Средний образ в каждом классе представлен вертикальной линией (см. правую часть рис. 5.2.2), отдельные образы в классе — точками (косые линии определяют разброс отдельных образов относительно среднего образа).

Первоначальный размер исходных образов —  $112 \times 92$  пикселя. При этом пространство признаков сокращено более чем в 100 раз.

Как видно из данного примера, метод РСА позволяет реализовать сокращение пространства признаков, однако не достаточно точно решить задачу кластеризации в пространстве признаков.

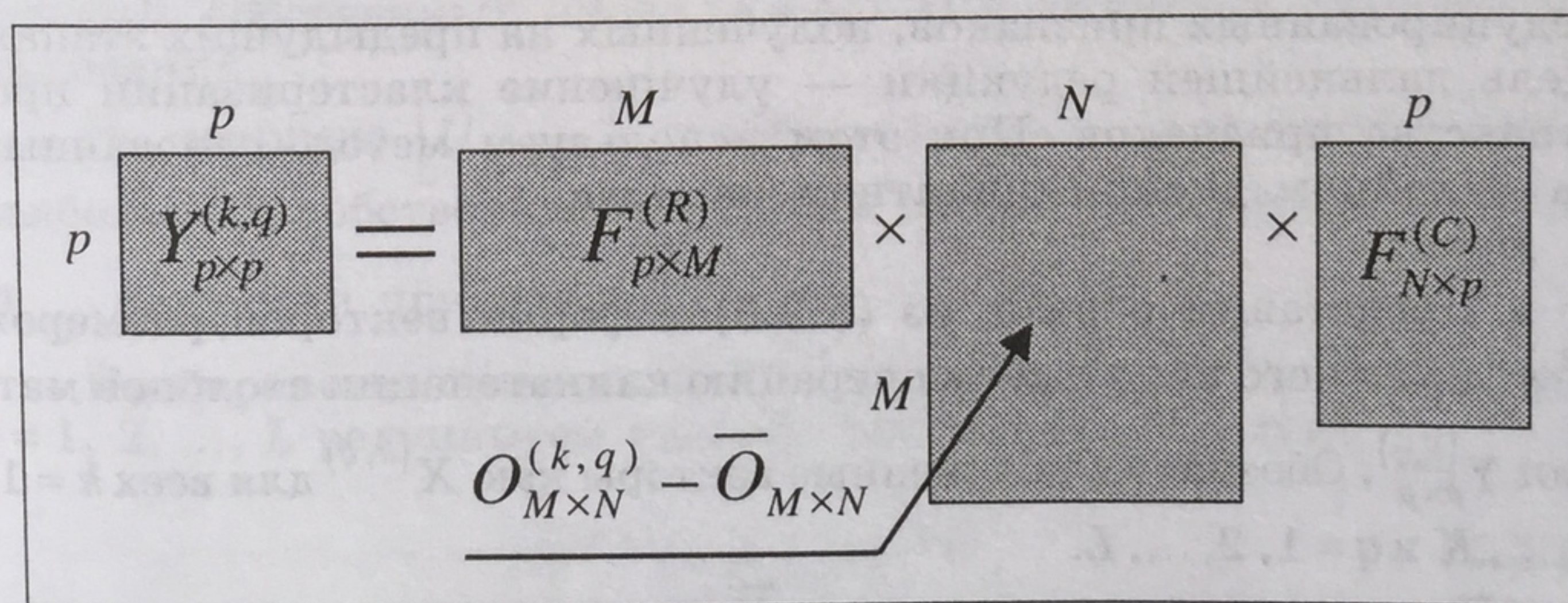


Рис. 5.2.1. Матричная схема редукции признаков для каждого образа в классе





Рис. 5.2.2. Отображение пространства редуцированных признаков

Здесь не максимизированы расстояния между классами (множествами признаков) и не минимизированы внутриклассовые расстояния. Иными словами, некоторые образы отходят от центров своих классов и так далеко, что попадают «на территорию» соседних классов (перекрывают другие классы), а сами центры классов имеют неравномерное распределение в пространстве признаков.

Улучшения кластеризации можно достичь при выполнении линейного дискриминантного анализа в редуцированном пространстве признаков, т. е. уже после процедуры KLT.

### 5.3. УМЕНЬШЕНИЕ ПРОСТРАНСТВА ПРИЗНАКОВ

Исходной информацией для нас теперь является пространство редуцированных признаков, полученных на предыдущих этапах. Цель дальнейшей редукции — улучшение кластеризации пространства признаков. При этом используем метод, основанный на линейном дискриминантном анализе.

Итак:

- Представим образы из (5.2.6) в форме векторов размером  $p^2 \times 1$ , для чего используем операцию канкатенации столбцов матриц  $Y_{p \times p}^{(k,q)}$ . Обозначим полученные векторы как  $X^{(k,q)}$  для всех  $k = 1, 2, \dots, K$  и  $q = 1, 2, \dots, L$ .

- Вычислим среднее значение  $\bar{X}$  всех признаков, а также среднее значение  $\bar{X}^{(k)}$  в каждом классе (для всех  $k = 1, 2, \dots, K$ ):



$$\left. \begin{aligned} \bar{X} &= \left( \sum_{k=1}^K \sum_{q=1}^L X^{(k,q)} \right) / (LK); \\ \bar{X}^{(k)} &= \left( \sum_{q=1}^L X^{(k,q)} \right) / L; \end{aligned} \right\} \quad k = 1, 2, \dots, K. \quad (5.3.1)$$

• Вычислим матрицы ковариаций  $W_{p^2 \times p^2}$  и  $B_{p^2 \times p^2}$ , из которых первая определяет ковариацию внутри классов, а вторая — между классами:

$$W_{p^2 \times p^2} = \sum_{k=1}^K \sum_{q=1}^L \left( X^{(k,q)} - \bar{X}^{(k)} \right) \cdot \left( X^{(k,q)} - \bar{X}^{(k)} \right)^T; \quad (5.3.2)$$

$$B_{p^2 \times p^2} = \sum_{k=1}^K \left( \bar{X}^{(k)} - \bar{X} \right) \cdot \left( \bar{X}^{(k)} - \bar{X} \right)^T. \quad (5.3.3)$$

• На основе (5.3.2) и (5.3.3) вычислим новую матрицу ковариации

$$H_{p^2 \times p^2} = W_{p^2 \times p^2}^{-1} \cdot B_{p^2 \times p^2}, \quad (5.3.4)$$

для которой определим собственные числа и соответствующие им собственные векторы:

$$D_{p^2 \times p^2} = \left[ U_{p^2 \times p^2} \right]^T H_{p^2 \times p^2} U_{p^2 \times p^2}, \quad (5.3.5)$$

где  $D_{p^2 \times p^2}$  — диагональная матрица собственных чисел;  $U_{p^2 \times p^2}$  — ортогональная матрица, строки которой определяют собственные векторы, соответствующие собственным числам.

• В диагональных элементах матрицы  $D_{p^2 \times p^2}$  найдем  $s$  наибольших собственных чисел ( $s = K - 1$ ) и запомним порядок их следования.

• Из матрицы  $\left[ U_{p^2 \times p^2} \right]^T$  выберем  $s$  строк, соответствующих  $s$  наибольшим собственным числам, и сформируем новую матрицу  $A_{s \times p^2}$  редукции признаков.

• Пространство признаков  $X^{(k,q)}$  для всех  $k = 1, 2, \dots, K$  и всех  $q = 1, 2, \dots, L$  редуцируем еще раз следующим образом:

$$\hat{X}^{(k,q)} = A_{s \times p^2} X^{(k,q)}, \quad (5.3.6)$$

где каждый вектор  $\hat{X}^{(k,q)}$  имеет размер  $s \times 1$ .



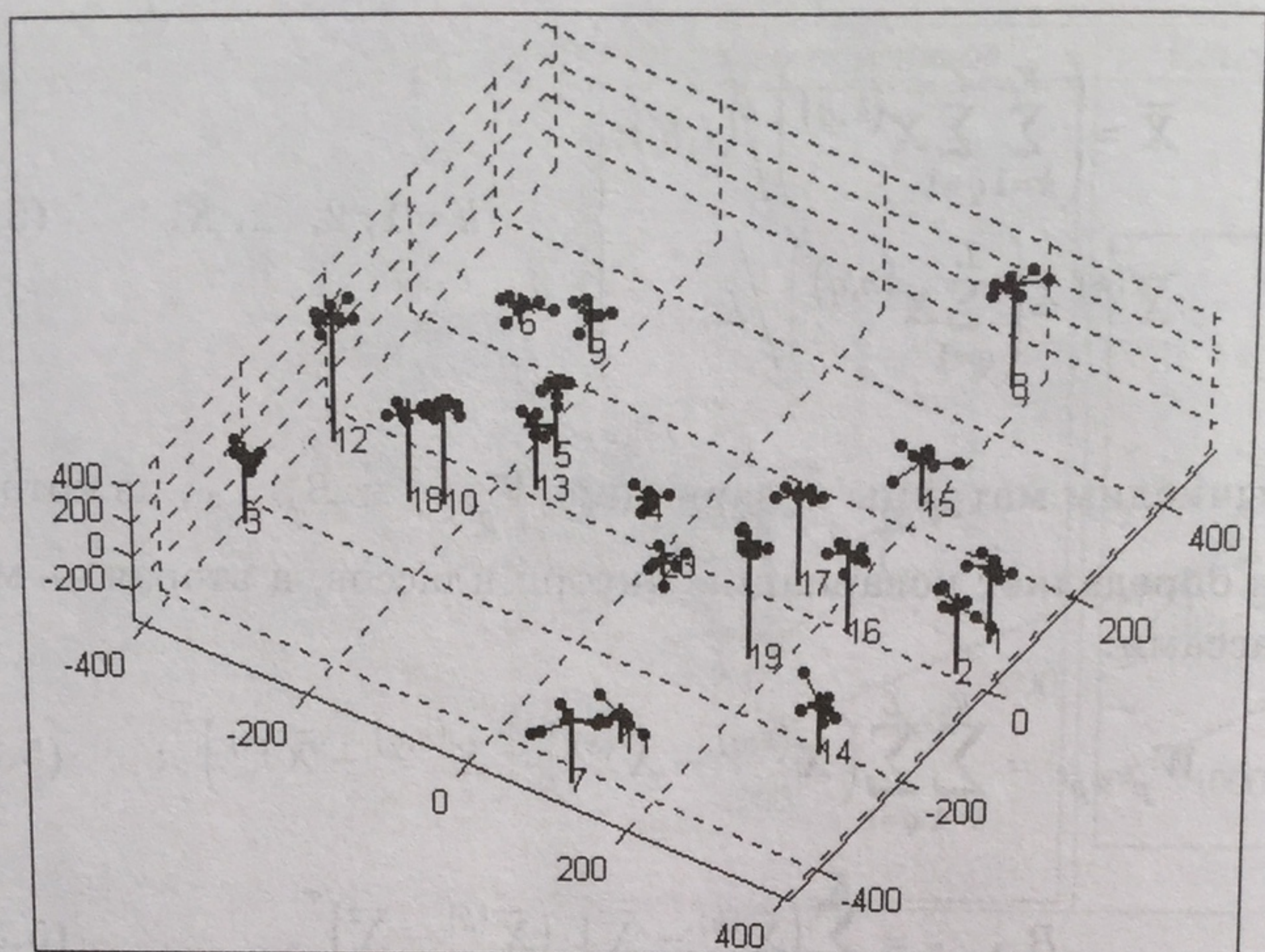


Рис. 5.3.1. Отображение 3D-пространства редуцированных признаков для метода PCA + LDA

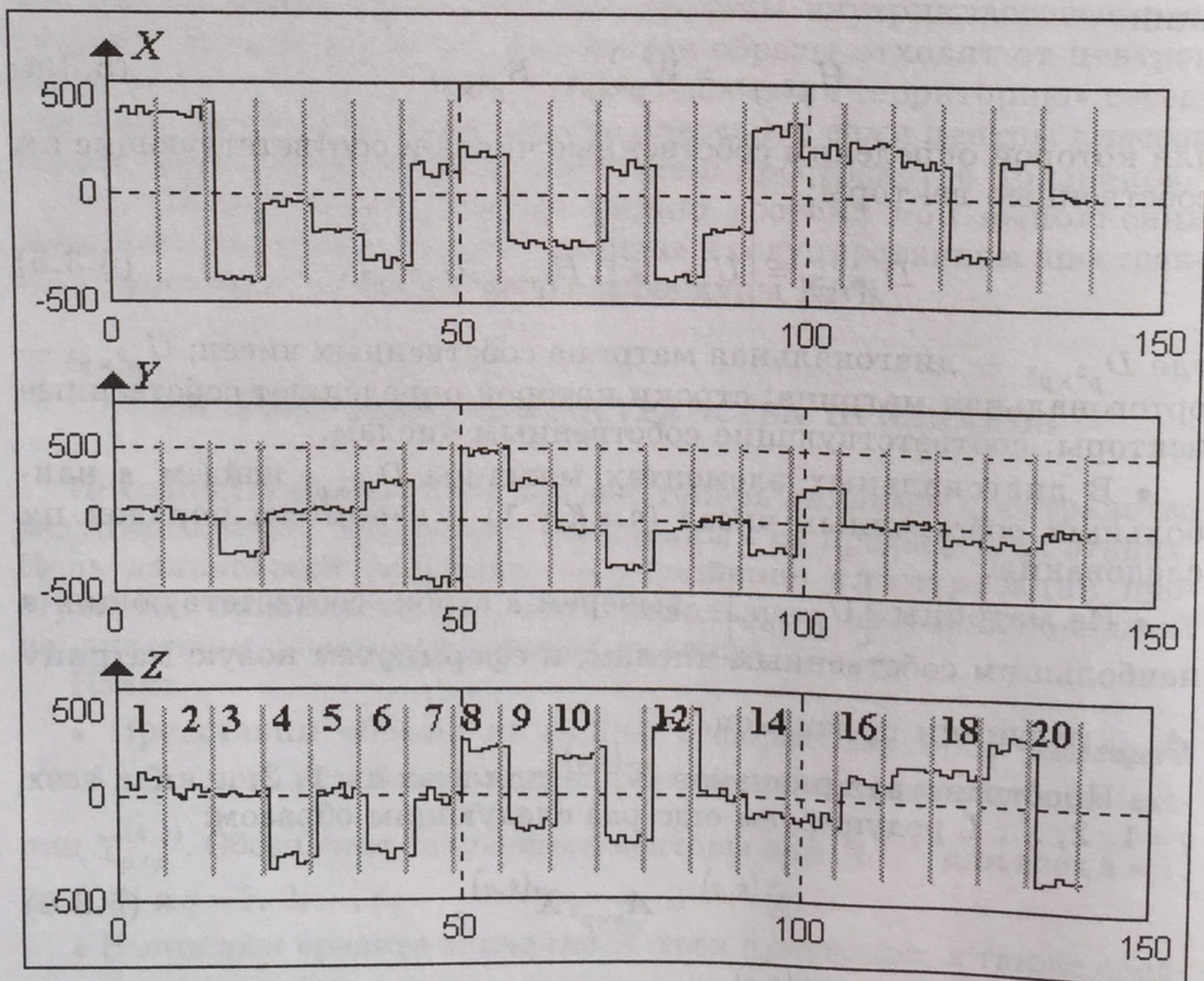


Рис. 5.3.2. Изменения первых трех признаков редуцированного пространства



Отображение редуцированного пространства признаков  $\hat{X}^{(k,q)}$  (для параметров ВБ:  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$ ) в системе координат 3D представлено на рис. 5.3.1. За координаты  $X$ ,  $Y$  и  $Z$  приняты первые три компоненты для каждого редуцированного вектора.

Как видно из представленного результата, применение метода LDA позволяет не только дополнительно сократить пространство признаков (общее сокращение составляет теперь более 500 раз), но и оптимизировать кластеризацию — середины классов лучше отдалены друг от друга (см. рис. 5.3.1), а «образы» внутри классов сильнее концентрируются у середин классов. Здесь уже практически не видно перекрытия кластеров и, следовательно, можно решить задачу распознавания простыми способами, используя, например, процедуру классификации по принципу минимального расстояния до середин классов.

Другой способ отображения редуцированного пространства признаков представлен на рис. 5.3.2: координатные кривые показывают изменение соответствующих признаков как внутри каждого класса (номера классов помечены цифрами от 1 до 20), так и всей базы данных. Здесь также можно оценить качество кластеризации и отметить классы, для которых есть риск неверного распознавания (см., например, признаки классов 5 и 13). При  $s > 3$  необходимо, очевидно, исследовать редуцированное пространство по всей совокупности признаков.

#### 5.4. ПРОЦЕСС РАСПОЗНАВАНИЯ КОНТРОЛЬНЫХ ОБРАЗОВ

Хотя известно, что параметр  $s$  выбирается из условий  $3 \leq s \leq K - 1$ , выбор конкретных (возможно, наилучших) параметров  $\{Q, L, p, s\}$  для заданного числа классов  $K$  возможен лишь при анализе синтезированной системы путем проверки способности распознавания  $(Q - L)$  последних образов в ВБ (т. е. образов  $O_{M \times N}^{(k,q)}$  с текущими индексами  $q = L + 1, L + 2, \dots, Q$ ).

Итак, пусть система распознавания синтезирована для конкретного значения параметров  $\{K, Q, L, p, s\}$ . В результате синтеза получены векторы  $\hat{X}^{(k,q)}$  редуцированного пространства призна-

ков, а также матрицы редукции  $F_{p \times M}^{(R)}, F_{N \times p}^{(C)}, A_{s \times p^2}$ .

Процесс распознавания нового образа  $O_{M \times N}^{(k,q)}$  (с текущими индексами  $q = L + 1, L + 2, \dots, Q$ ) реализуется в три этапа.

1. Центрирования текущего образа относительно среднего образа  $\bar{O}_{M \times N}$  (см. (5.2.1)).

2. Редукции признаков для центрированного образа с использованием матриц  $F_{p \times M}^{(R)}, F_{N \times p}^{(C)}, A_{s \times p^2}$  и получением вектора  $\tilde{X}$ .



3. Вычисление расстояний между вектором  $\tilde{X}$  и векторами средних  $\tilde{X}^{(k)}$  всех классов и определение номера класса, которому соответствует минимум расстояния.

Реализацию первых двух этапов запишем в следующей форме:  
**редукция I**

$$\tilde{Y}_{p \times p}^{(k,q)} = F_{p \times M}^{(R)} \left( O_{M \times N}^{(k,q)} - \bar{O}_{M \times N} \right) F_{N \times p}^{(C)} \quad (5.4.1)$$

для всех  $k = 1, 2, \dots, K$  и всех  $q = L + 1, L + 2, \dots, Q$ ;

**редукция II**

$$\tilde{X} = A_{s \times p^2} \tilde{X}, \quad (5.4.2)$$

где вектором  $\tilde{X}$  размера  $p^2 \times 1$  обозначим результат конкатенации столбцов матрицы  $\tilde{Y}_{p \times p}^{(k,q)}$ .

Расстояния между векторами определим по одной из известных метрик, например, в метрике Евклида, в форме

$$d_k = \sqrt{(\tilde{x}_1 - \bar{x}_1^{(k)})^2 + (\tilde{x}_2 - \bar{x}_2^{(k)})^2 + \dots + (\tilde{x}_s - \bar{x}_s^{(k)})^2} \quad (5.4.3)$$

для всех  $k = 1, 2, \dots, K$ .

Графически расстояние от распознаваемого образа до ближайшего класса (расстояние по трем первым координатам — трем

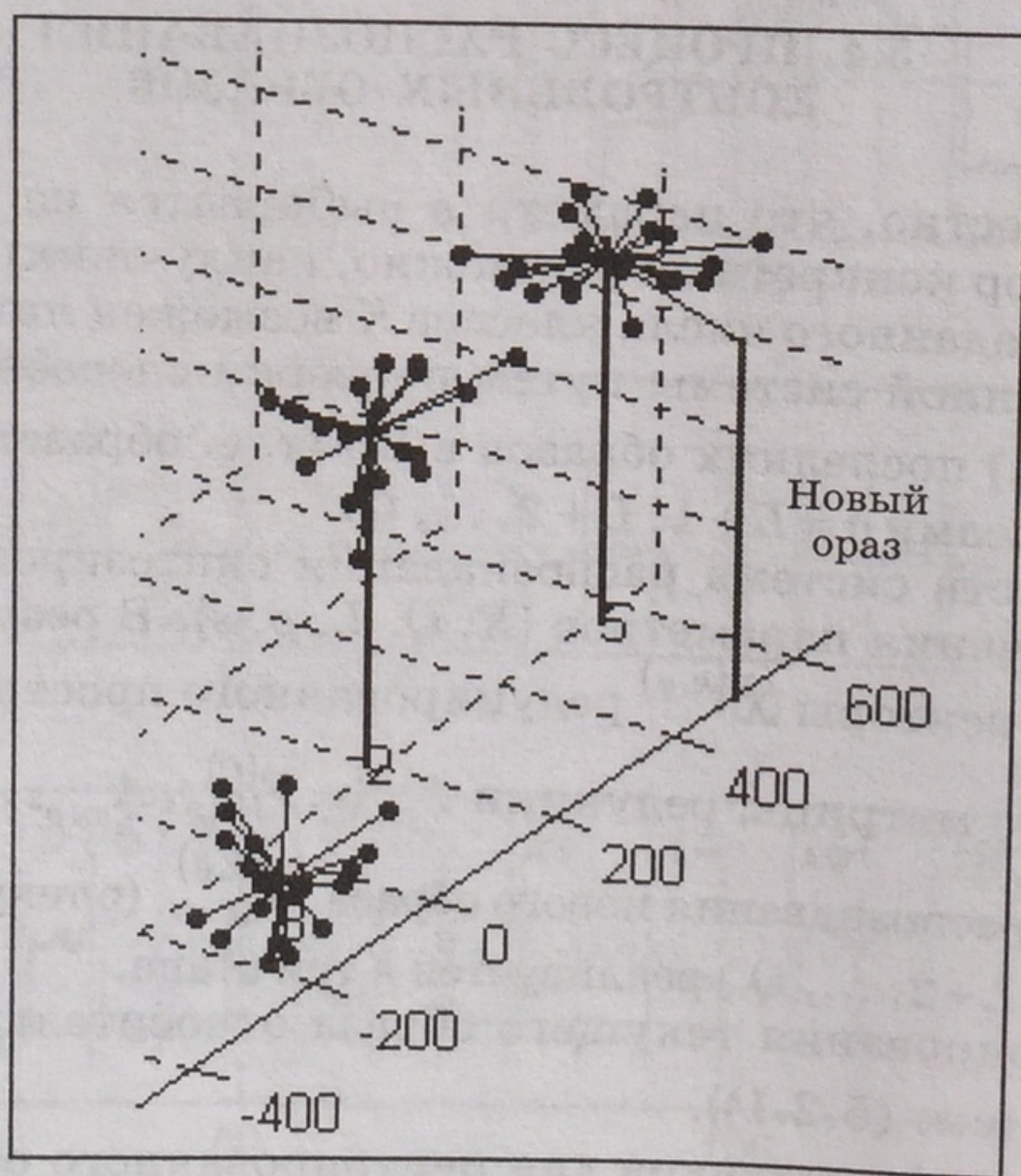


Рис. 5.4.1. Графическое представление расстояния в пространстве признаков



первым компонентам векторов  $\bar{X}^{(k)}$  и  $\tilde{X}$  в 3D-пространстве редукцированных признаков показано на рис. 5.4.1. Здесь цифрами 2, 5 обозначены высоты средних значений классов (среднее значение третьей компоненты векторов  $X^{(k)}$ ). Отдельной вертикальной линией обозначен распознаваемый образ. Линия, соединяющая вершины нового образа и середины 5-го класса, — оцениваемое расстояние. Здесь ближайшим соседом является 5-й класс.

## 5.5. ВЫБОР ПАРАМЕТРОВ СИСТЕМЫ РАСПОЗНАВАНИЯ

Оценить возможности распознавания представленного выше метода можно было бы в аналитической форме, следуя, например, [Tsa98, Swe96, Swe99]. Однако напомним, что представленный метод складывается из четырех обоснованных, проверенных и, по мнению авторов, наилучших подходов, используемых при решении проблем распознавания образов, подходов, высокое качество которых было убедительно показано в разд. 3.7.

В дополнение к этим доказательствам представим результаты некоторых экспериментов по распознаванию контрольных образов ВБ (т. е. образов  $O_{M \times N}^{(k, q)}$ , для  $q = L + 1, L + 2, \dots, Q$ ), полученные для различных параметров  $\{K, Q, L, p, s\}$ , а также этих же образов, но подвергнутых искажениям — масштабированию/расфокусировке, наложению шума и др.

### 5.5.1. ИЗМЕНЕНИЕ ЧИСЛА КЛАССОВ И ЧИСЛА ОБРАЗОВ В КЛАССЕ

Цель настоящего эксперимента — оценить возможности синтеза системы распознавания и качества распознавания при смене параметра  $K$ . При этом примем параметр  $K = 4, 20, 40$  и  $49$ .

#### 1. Число классов $K = 4$ , число образов $L = 6$ .

Исходная база данных — выбранные четыре класса образов из базы данных [Orl\*\*]. Каждый образ имеет размеры  $112 \times 92$  пикселя.

Оценка эффекта распознавания для такого малого числа классов показала, что минимальное число образов в каждом классе не должно быть меньше 6, поскольку матрицы ковариации, используемые в методах PCA и LDA, при  $L < 6$  становятся особенными и в таком случае нельзя гарантировать стабильность редукции исходного признакового пространства. Один из вариантов виртуальной базы данных для случая  $K = 4$  и  $L = 6$  представлен на рис. 5.5.1.

Собственные числа, вычисленные для матриц  $R_{M \times M}$  и  $C_{N \times N}$ , упорядоченные по убыванию, приведены на рис. 5.5.2 (в левой части).





Рис. 5.5.1. База данных для параметров  $K = 4$ ,  $Q = 10$  и  $L = 6$

Здесь видно, что параметр  $p$  не может превосходить значения 10. В данном случае редукция I (метод PCA) реализована при  $p = 4$ . При этом, поскольку число классов  $K = 4$ , то  $s = 3$ . Пространство редуцированных признаков показано на рис. 5.5.2 (правая сторона).

Пространство редуцированных признаков в 3D-форме, полученное при редукции I и II, показано на рис. 5.5.3. Во-первых, виден эффект дискриминантного анализа, а во-вторых, пространство признаков сократилось здесь на величину, близкую к 3500, поскольку  $112 \cdot 92 / 3 \approx 3500$ .

Результаты распознавания четырех контрольных образов в каждом классе представлены на рис. 5.5.4, где с левой стороны в

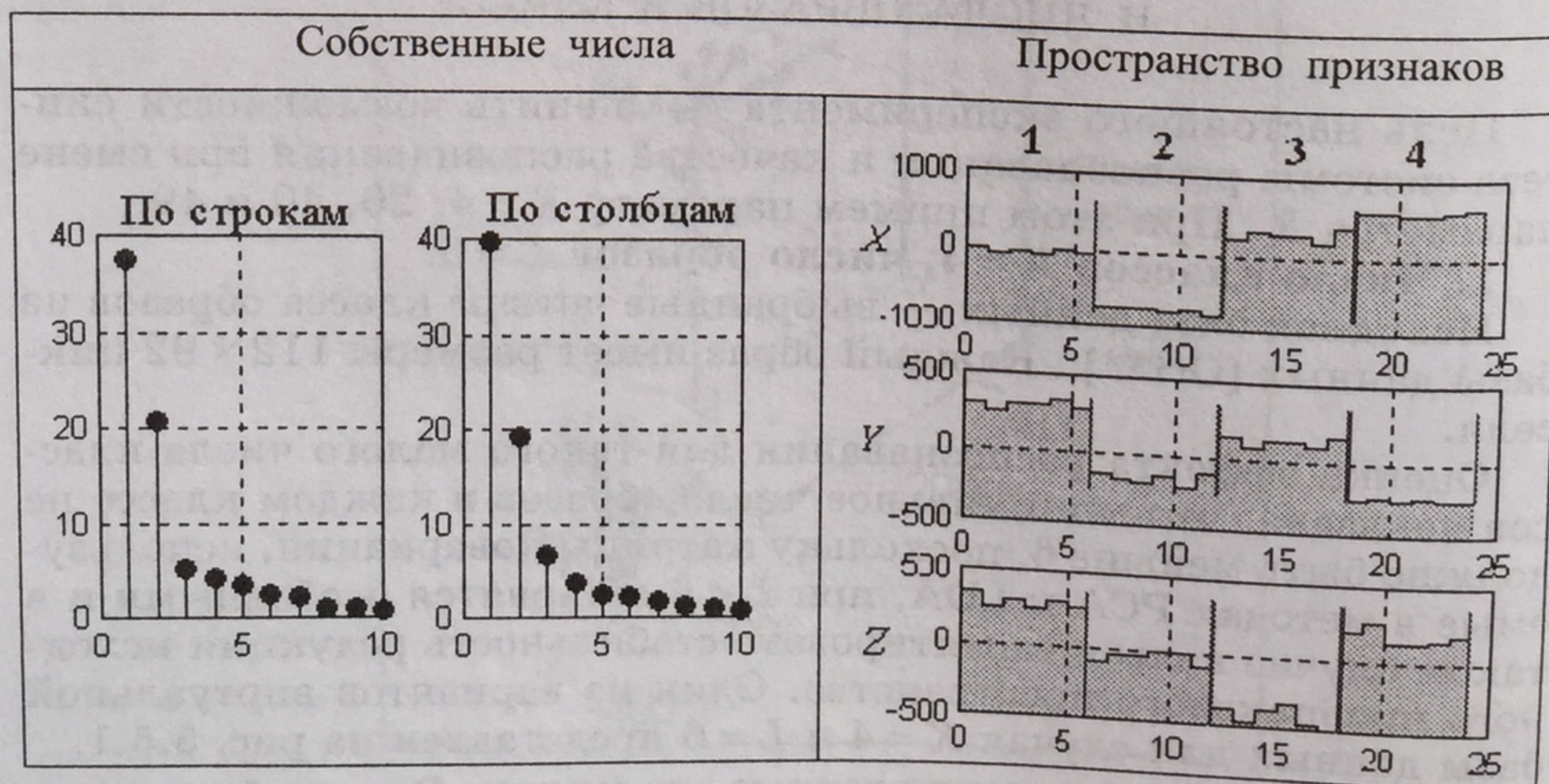


Рис. 5.5.2. Собственные числа и пространство редуцированных признаков ( $p = 4$ ,  $s = 3$ )



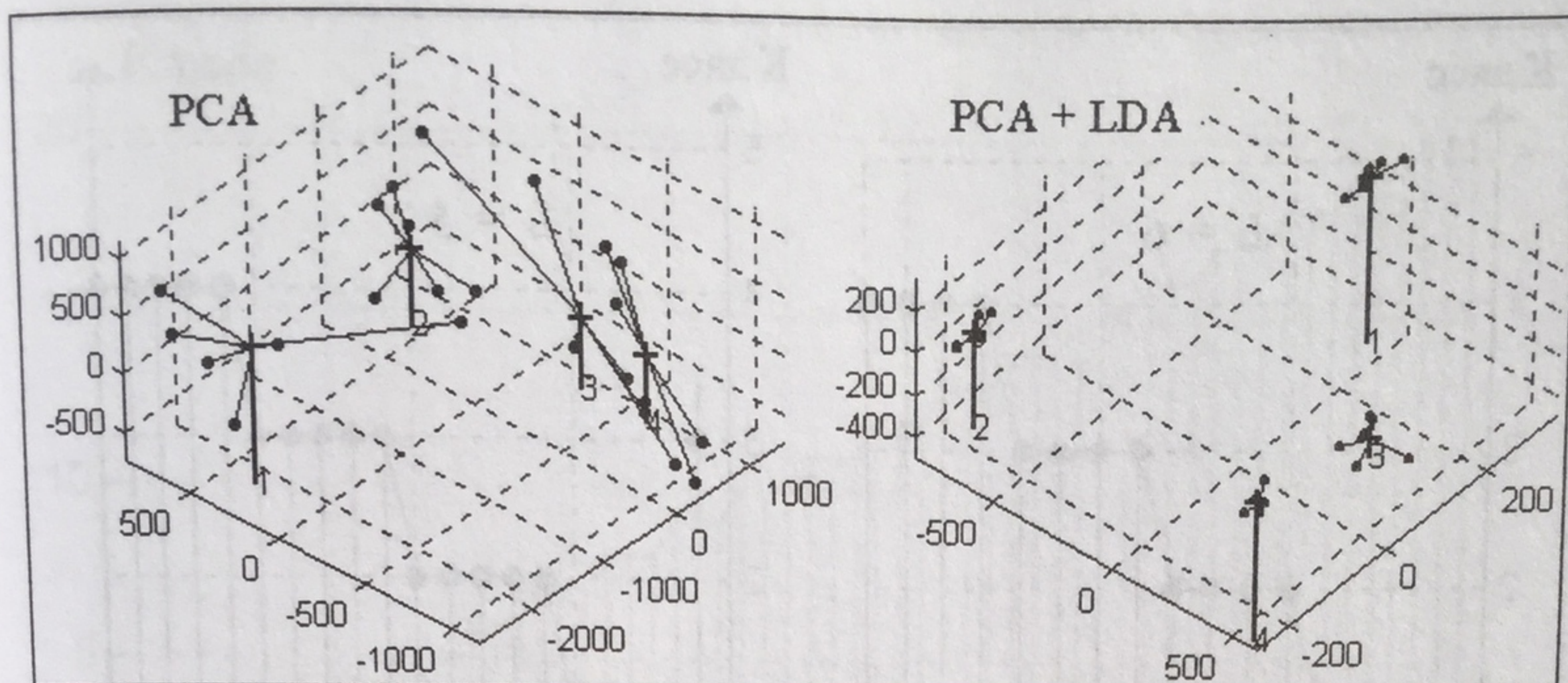


Рис. 5.5.3. Пространство редуцированных признаков в 3D-форме

каждом столбце (каждом выделенном классе) показаны контрольные образы, а с правой стороны — наиболее «близкие образы», найденные в ВБ. Обратим внимание на то, что наклон головы и поза (соответствующая углу/точке наблюдения за лицом), а также выражение лица (гримаса) подобранных «близких образов» практически соответствуют друг другу. Зеркальная симметрия, наблюдаемая в некоторых парах, — результат отсутствия в ВБ соответствующего образа.

Отмеченное «соответствие» объясняется тем, что первая и вторая компоненты редуцированных признаков «отвечают» за поворот и позу головы, а третья — собственно за выражение лица.

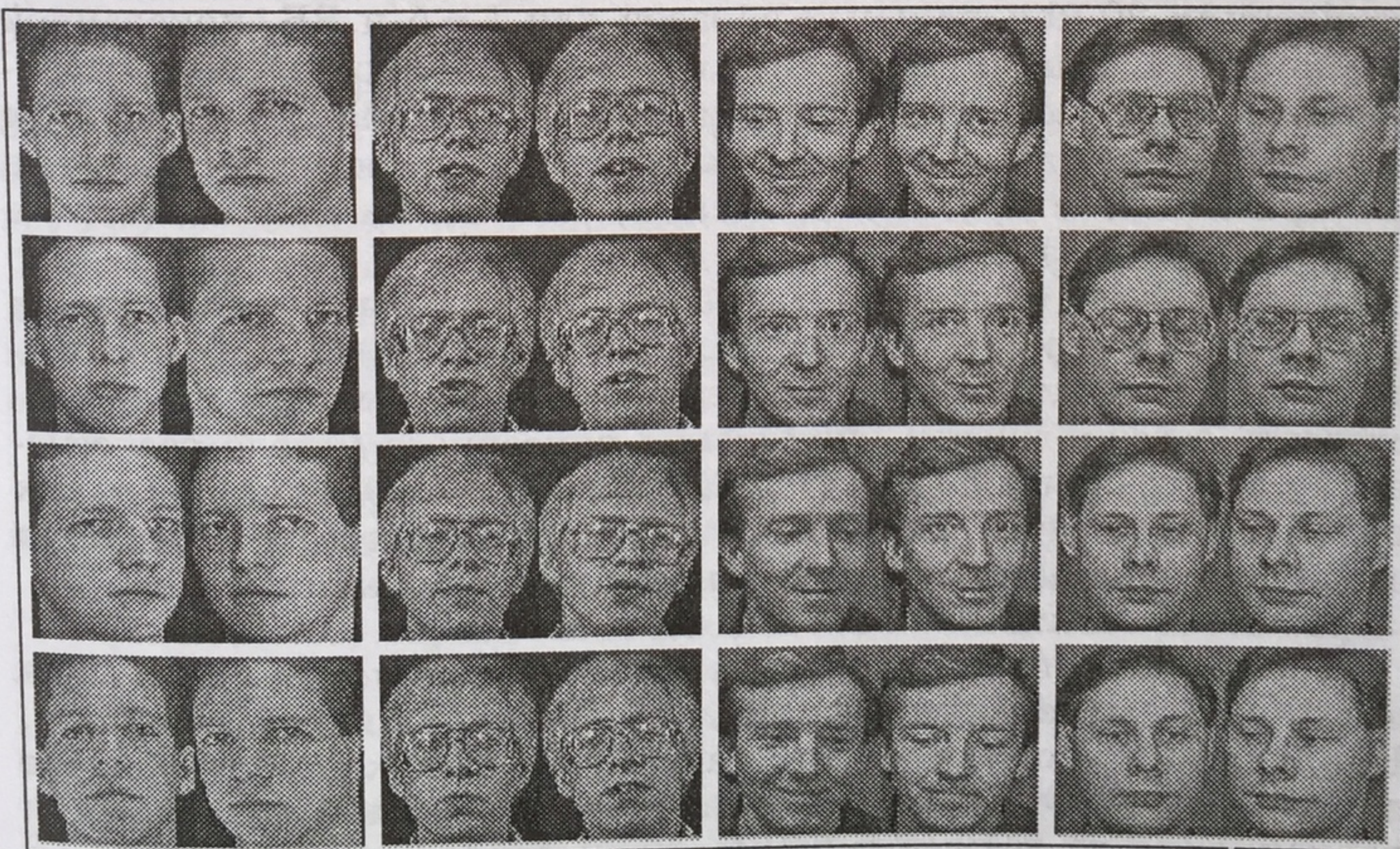


Рис. 5.5.4. Подбор близких образов из ВБ в результате распознавания 16 контрольных образов



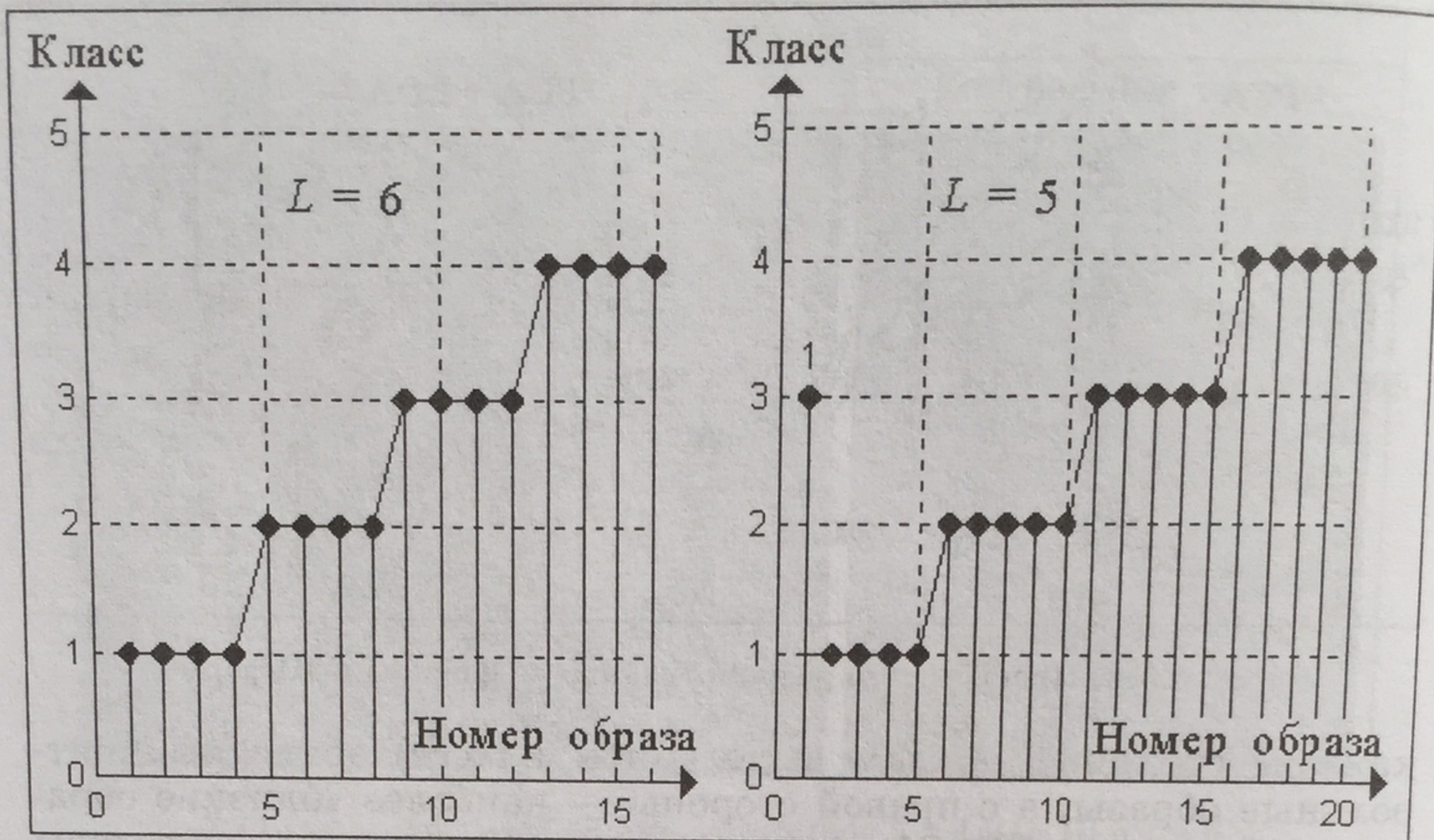


Рис. 5.5.5. Итоговые результаты распознавания контрольных образов для системы с параметрами  $K = 4$ ,  $Q = 10$ ,  $p = 4$ ,  $s = 3$

При этом влияние компоненты в процессе распознавания (в результате выбора близкого образа) тем выше, чем ниже порядковый номер компоненты.

В итоге распознавание 16 контрольных образов (рис. 5.5.5, левый график) составило 100 %.

С правой стороны на рис. 5.5.5 графически показан итог распознавания 20 контрольных образов для  $L = 5$  и ВБ, показанной на рис. 5.5.1. Как видно, здесь уже имеется одно неверное распознавание в классе номер 1 (при общем итоге 95 %). При этом следует подчеркнуть, что образы класса 1 имеют значительные (по сравнению с другими классами) изменения как в повороте головы, так и размеров.

Аналогичные эксперименты распознавания были выполнены и для других четырех выбранных классов базы [Orl\*\*]. При этом везде при  $L = 6$  распознавание составило 100 %, что может свидетельствовать о возможности применения описанного метода распознавания для граничных (снизу) параметров  $K$  и  $L$ . Обычно при таких параметрах системы распознавания синтезируются при решении задач контроля доступа ограниченного числа пользователей.

## 2. Число классов $K = 20, 40, 49$ .

Увеличение числа классов  $K$  требует и увеличения числа образов  $L$  в ВБ, и увеличения параметра  $p$  в границах возможного. Наилучший выбор последнего параметра возможен только в ходе анализа качества распознавания синтезированной системы при вариациях  $p$  во всем возможном диапазоне.

Представим результаты работы системы с параметрами  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$ . Образы исходной базы данных и результат



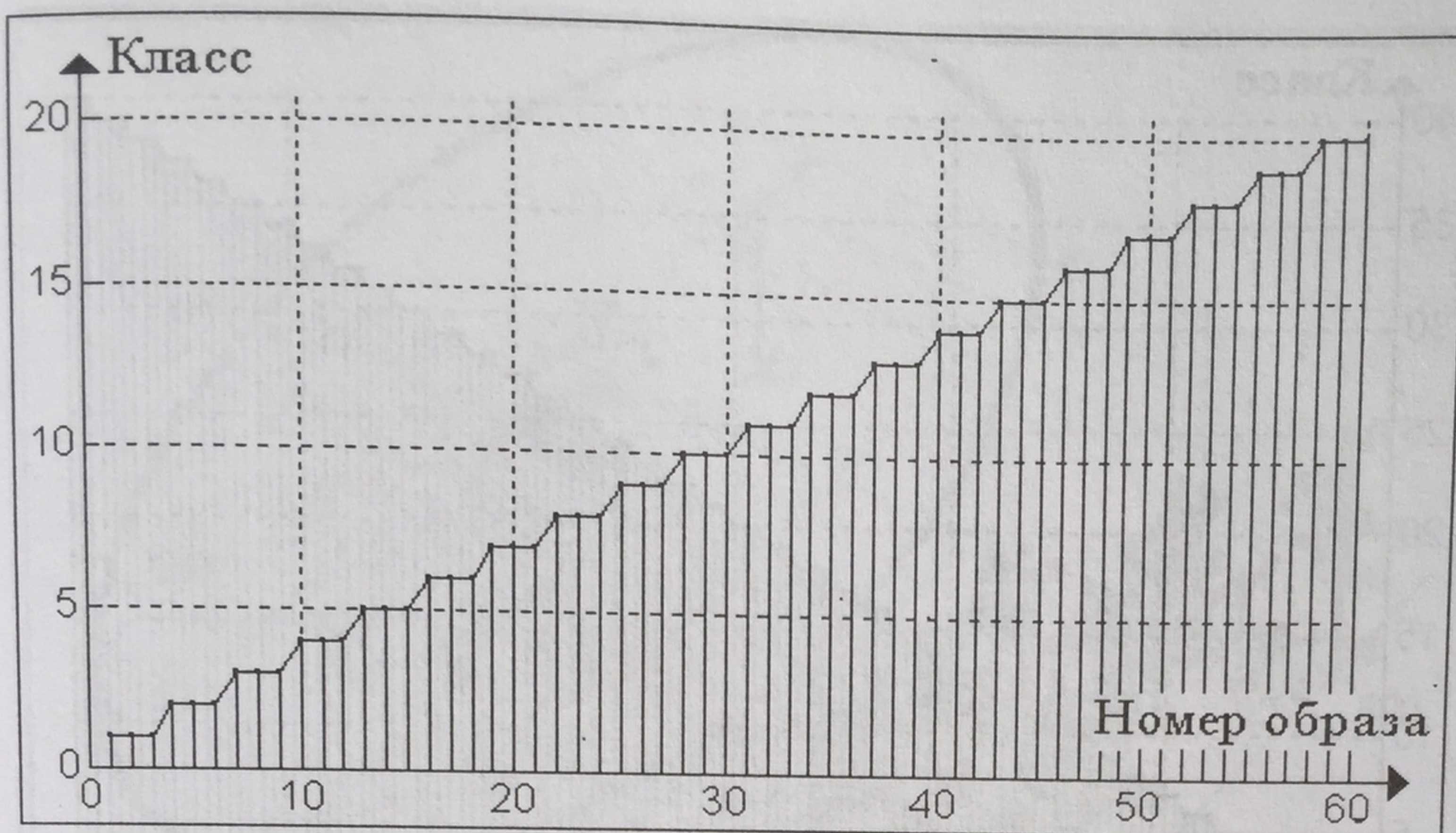


Рис. 5.5.6. Итоговые результаты распознавания 60 контрольных образов для системы с параметрами  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$

построения виртуальной базы для одного из классов базы [Orl\*\*] были приведены в предыдущем разделе, а пространство редуцированных признаков (в 2D- и 3D-форме) для этого случая показано также в предыдущем разделе на рис. 5.3.1 и 5.3.2.

Результаты распознавания 60 образов (по 3 образа в каждом классе) для системы с параметрами  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$  представлены на рис. 5.5.6. Очевидно, что распознавание контрольных образов достигло 100 %.

Для системы распознавания с параметрами  $K = 40$ ;  $L = 7$ ;  $p = 7$ , 8, 9, 10 и  $s = 39$  итоговые результаты распознавания контрольных образов представлены на рис. 5.5.7. Здесь также распознаются все 100 % контрольных образов.

Все результаты, представленные выше, основаны на базе данных [Orl\*\*]. Не исключено, что высокие показатели распознавания (поиска образа в базе, который соответствует контрольному образу) связаны с качеством самой базы [Orl\*\*].

Отметим особенности данных этой базы:

- высокое разрешение каждого изображения при относительно малых его размерах;
- относительно одинаковый размер собственно лиц;
- размещение образа лица практически на поле всего изображения;
- практически неизменный темный фон (текстура) и одинаковое освещение;
- незначительные изменения в поворотах головы в плоскости XY изображения, оси Z и оси X (пункта осмотра головы).

Интересно проверить представленный выше метод (качество распознавания/поиска в базе) в случае, когда исходная база данных



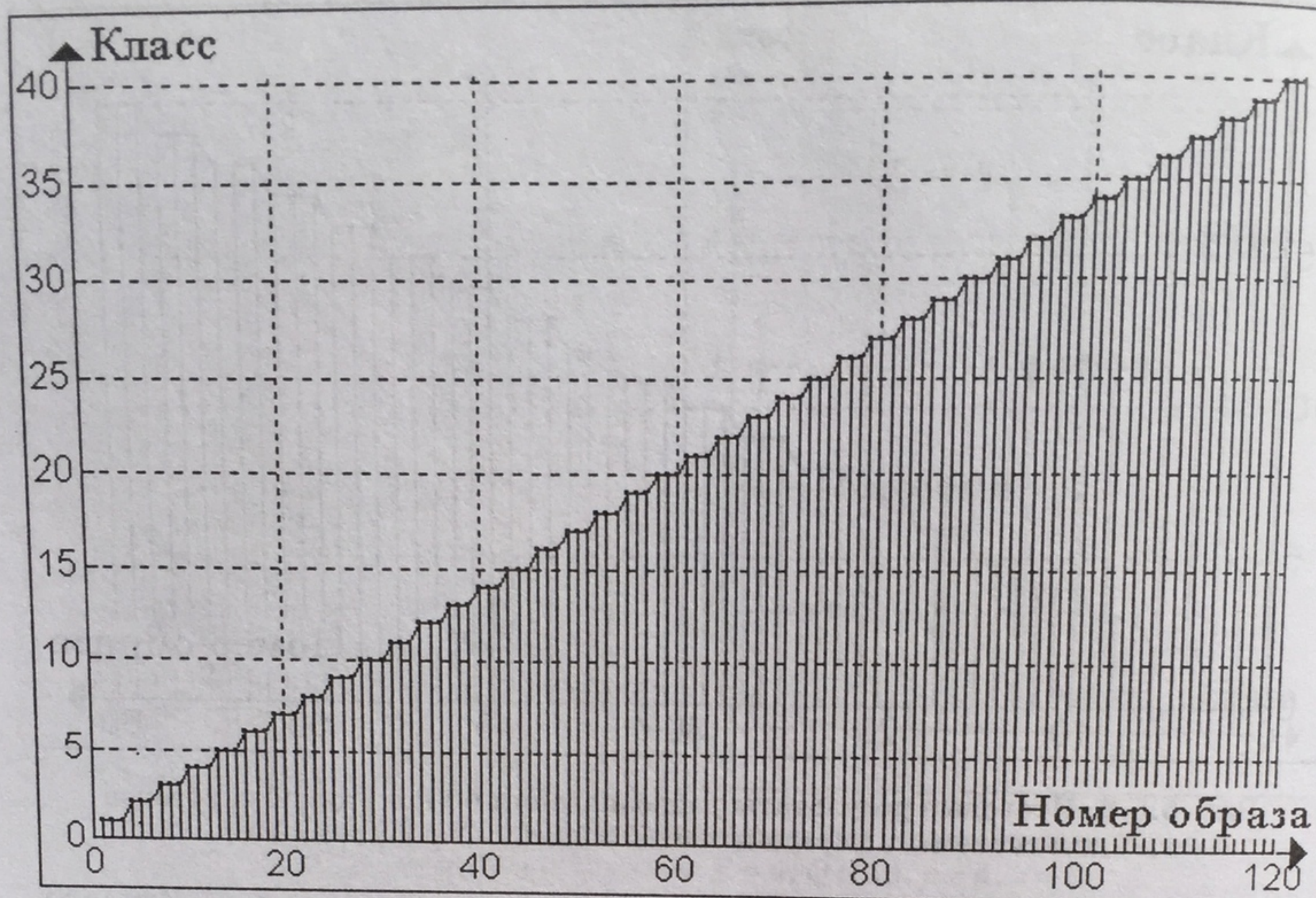


Рис. 5.5.7. Итоговые результаты распознавания 120 контрольных образов для системы с параметрами  $K = 40$ ;  $L = 7$ ;  $p = 7, 8, 9, 10$  и  $s = 39$



Рис. 5.5.8. Классы 41 – 46 новой базы данных



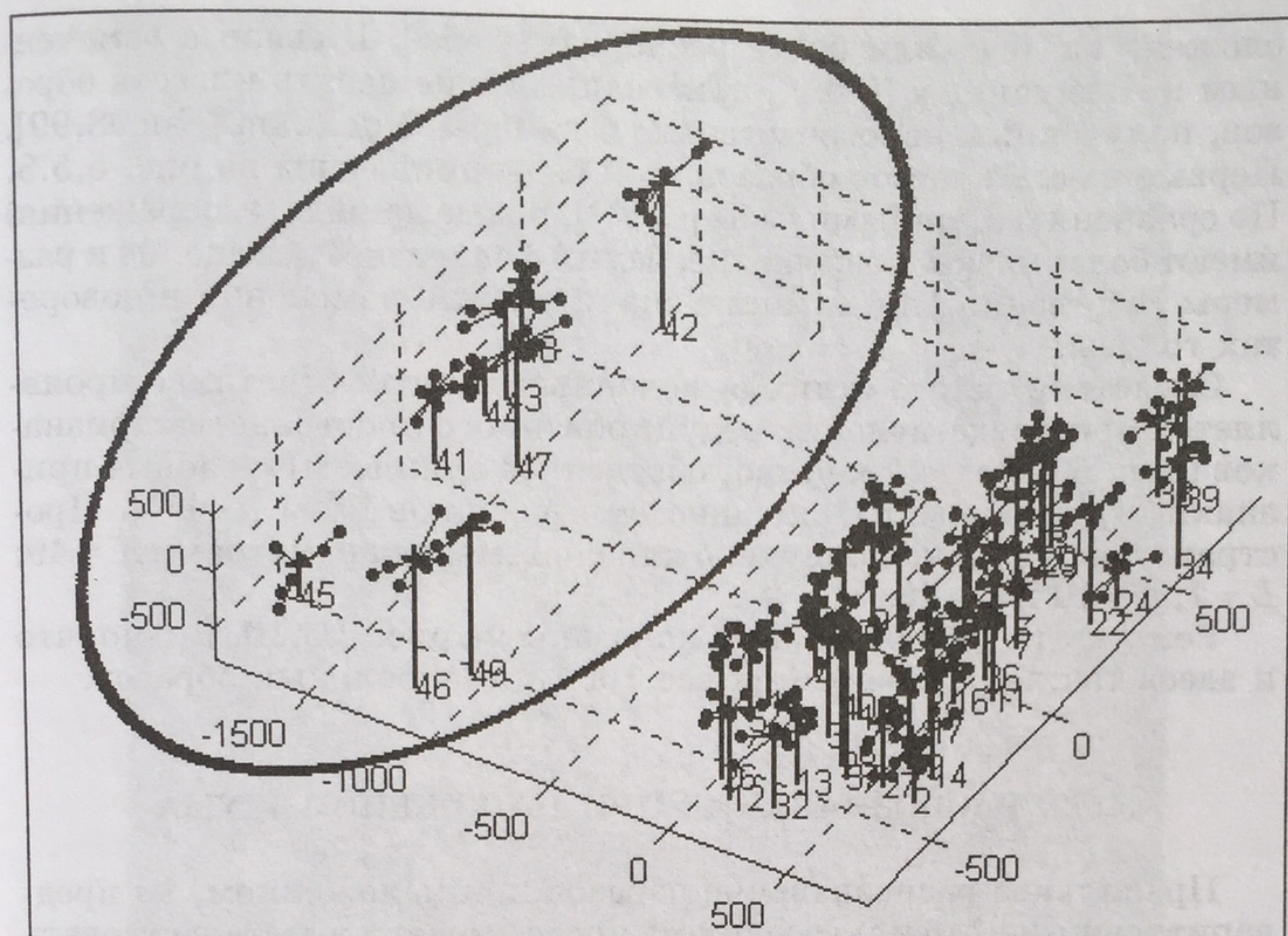


Рис. 5.5.9. Пространство редуцированных признаков для системы с параметрами  $K = 49$ ;  $L = 7$ ;  $p = 10$  и  $s = 48$

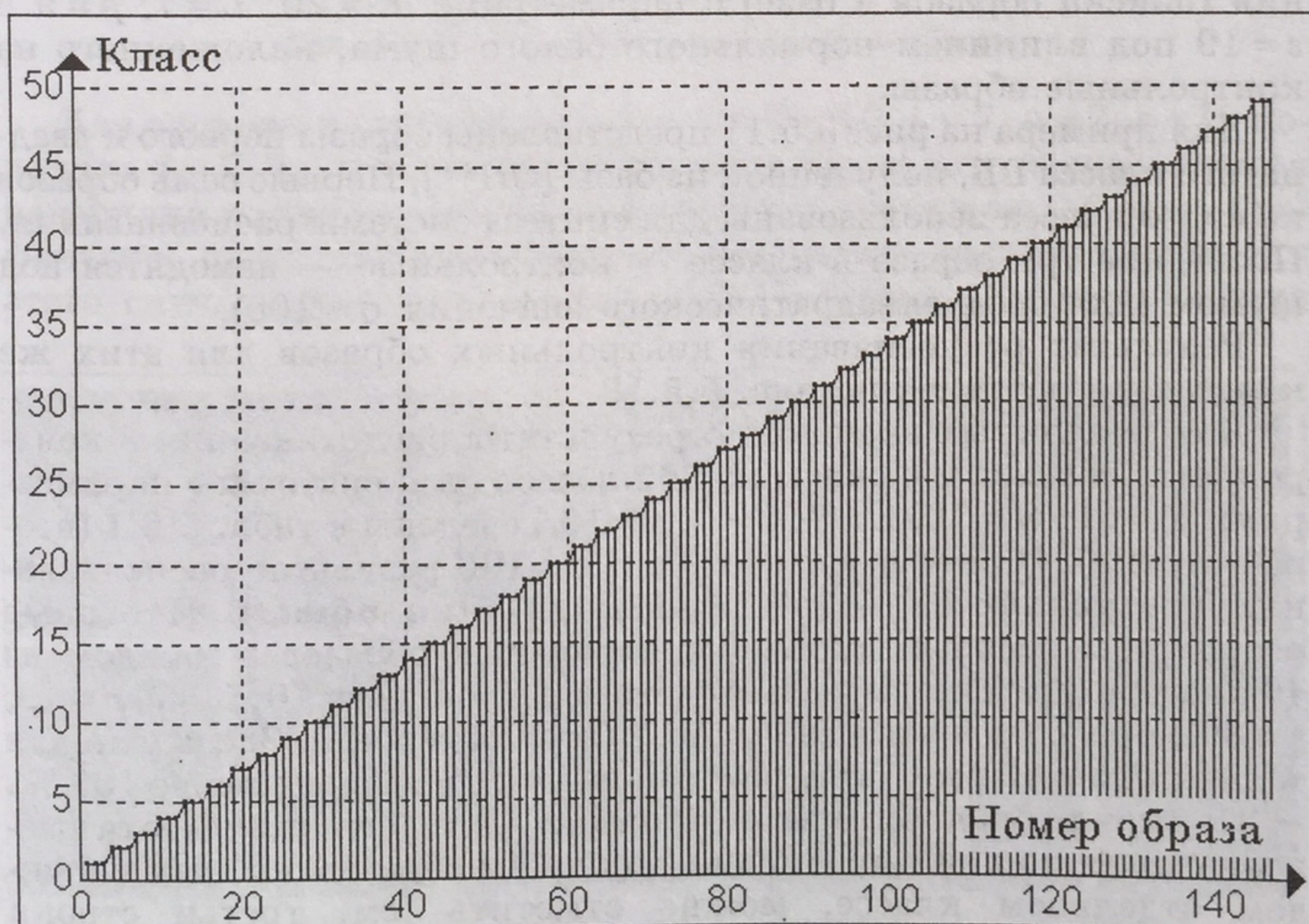


Рис. 5.5.10. Итоговые результаты распознавания 147 контрольных образов для системы распознавания с параметрами  $K = 49$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 10$ ,  $s = 48$



сложена из двух или более разнородных баз. В связи с отмеченным к базе данных [Orl\*\*] присоединим еще девять классов образов, полученных непосредственно с камеры QuickCam [Kuc98,99]. Первые шесть классов образов этой базы приведены на рис. 5.5.8. По сравнению с данными базы [Orl\*\*], новые данные (изображения) имеют более низкое разрешение, белый фон, разное освещение и размеры собственно лиц, а также значительные изменения в поворотах головы.

Следствие такого отличия исходных данных отчетливо проявляется при представлении редуцированного пространства признаков (рис. 5.5.9 пространство, обведенное эллипсом) — новые признаки сгруппированы отдельно от признаков базы [Orl\*\*]. Пространство признаков получено для системы с параметрами  $K = 49$ ;  $L = 7$ ;  $p = 10$  и  $s = 48$ .

Результаты распознавания показаны на рис. 5.5.10. Видно, что и здесь также распознаются все 100 % контрольных образов.

#### 5.5.2. РАСПОЗНАВАНИЕ ПРИ НАЛОЖЕНИИ ШУМА

Правильное распознавание образов при наложенном, но предварительно не отфильтрованном шуме может свидетельствовать о робастности системы распознавания к изменениям первичных признаков. Представим результаты работы системы распознавания (поиска образов в базе) с параметрами  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$  под влиянием нормального белого шума, наложенного на контрольные образы.

Для примера на рис. 5.5.11 представлены образы первого и двадцатого класса ВБ, полученной из базы [Orl\*\*]. Первые семь образов каждого класса использованы для синтеза системы распознавания. Последние три образа в классе — контрольные — находятся под шумом (для среднеквадратического значения  $\sigma = 100$ ).

Результат распознавания контрольных образов для этих же классов представлен на рис. 5.5.12.

Для различных значений  $\sigma$  результаты распознавания 3 контрольных образов в каждом из 20 классов (для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$ ) сведены в табл. 5.5.1 (вторая строка). Из нее видно, что для  $\sigma = 130$  результат распознавания составляет 98 % (не распознан *один образ*). Итоговые результаты распознавания 3 контрольных образов в каждом из 40 классов (для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$ ) представлены в табл. 5.5.2 (вторая строка). Здесь уже для значений  $\sigma \geq 70$  результат распознавания составляет менее 100 %.

Применяя мажоритарный принцип — оценку результата распознавания по двум правильно узнанным образам из трех в каждом отдельном классе, можно отметить (см. третьи строки табл. 5.5.1 и 5.5.2), что в обоих случаях распознавание прошло успешно со 100 % -ным результатом для значений  $\sigma \geq 110$ .

Д  
лож  
наве  
фиче  
этого

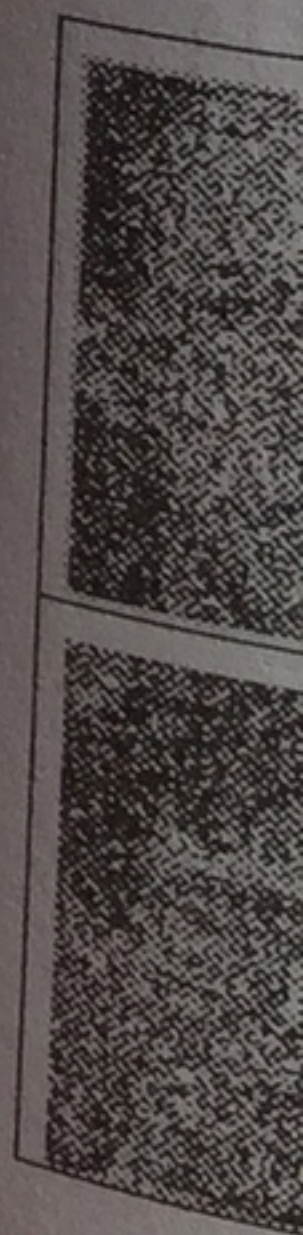


Рис. 5.



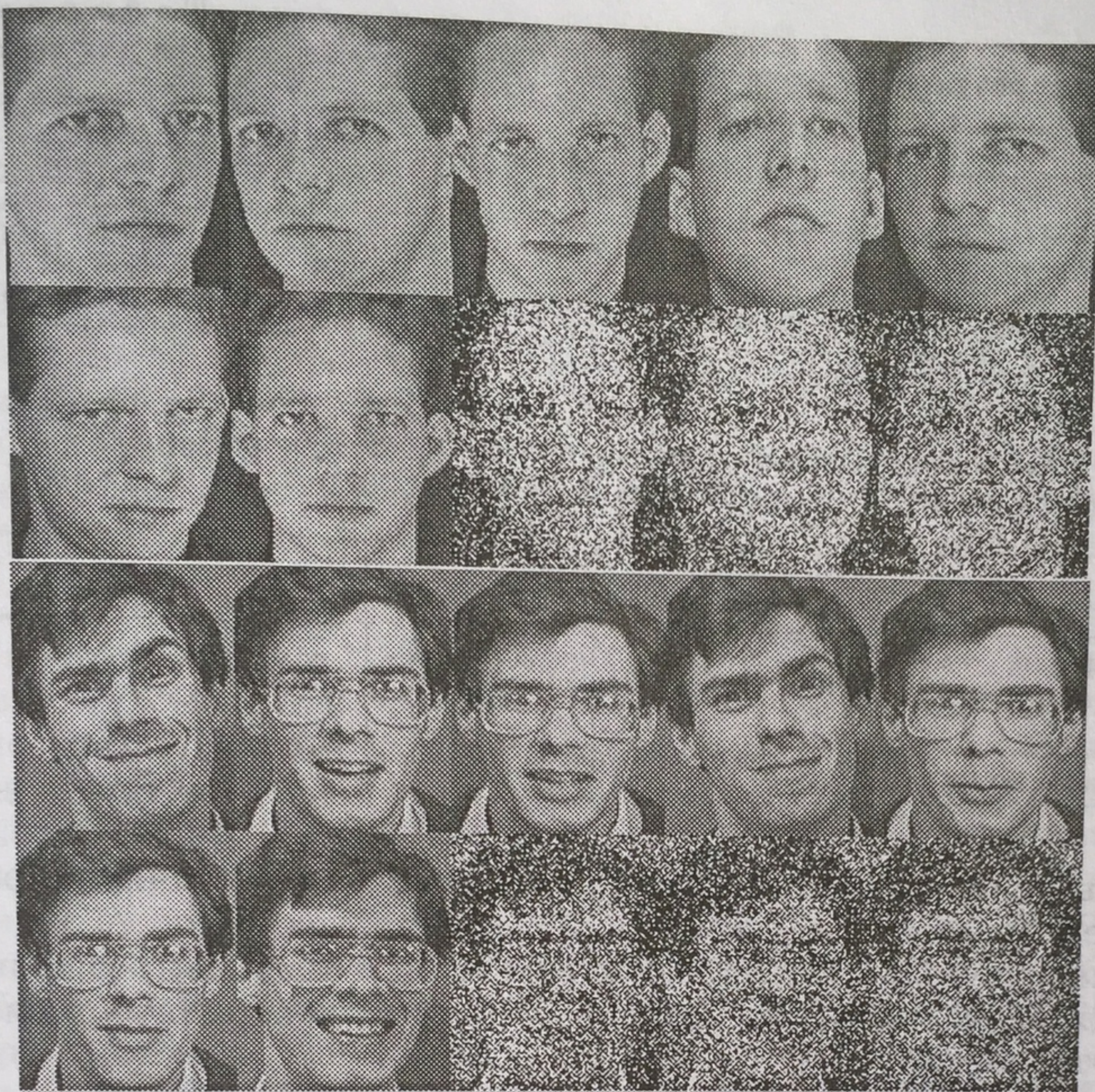


Рис. 5.5.11. Исходные и контрольные образы

Для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  положим  $\sigma = 150$  с тем, чтобы в итоговом результате распознавания наверняка появились неверно классифицированные образы. Графическое представление итогового результата распознавания для этого случая приведено на рис. 5.5.13 (левая сторона). Оценка

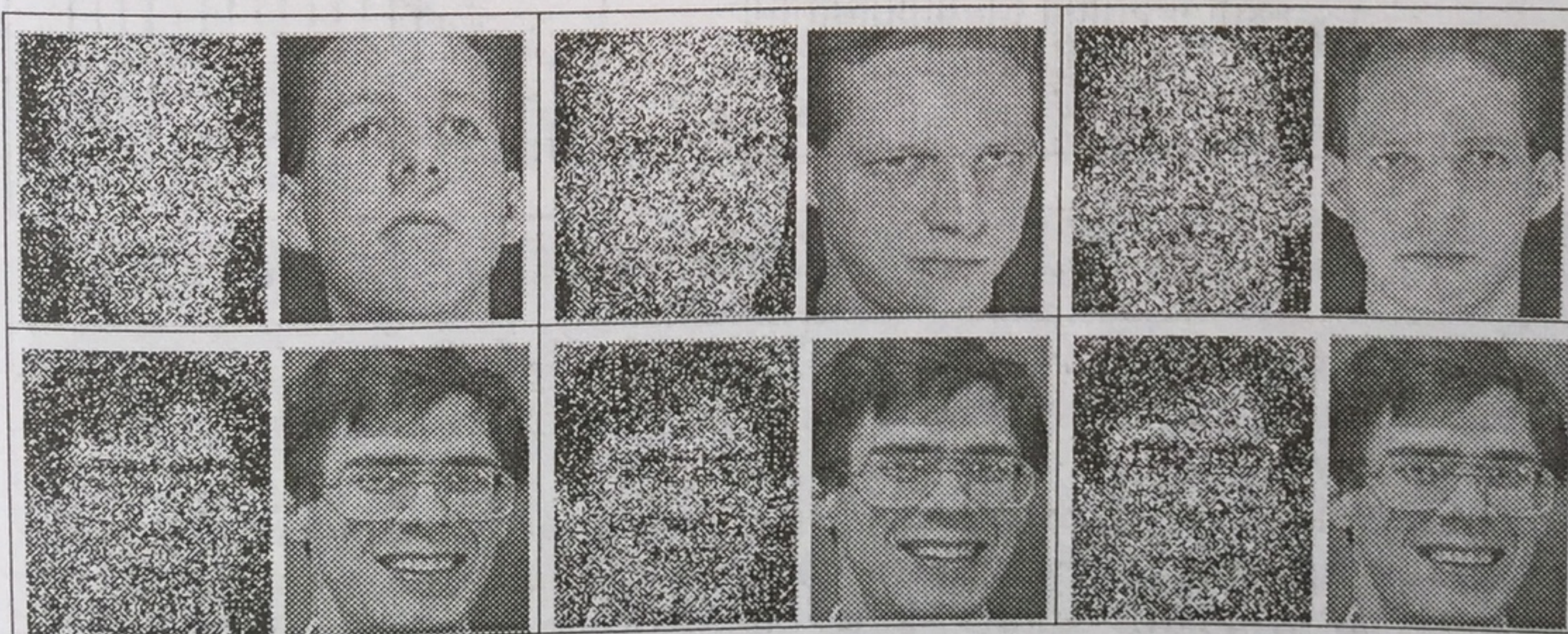


Рис. 5.5.12. Некоторые результаты распознавания зашумленных контрольных образов системой с параметрами  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$



Т а б л и ц а 5.5.1

Итоговые результаты распознавания 60 контрольных образов для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$

| $\sigma$                     | 50  | 60  | 70  | 80  | 90  | 100 | 110 | 120 | 130 |
|------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 60 образов, % распознавания  | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 98  |
| Два из трех, % распознавания | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Т а б л и ц а 5.5.2

Итоговые результаты распознавания 120 контрольных образов для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$

| $\sigma$                     | 50  | 60  | 70   | 80   | 90   | 100  | 110  | 120  | 130  |
|------------------------------|-----|-----|------|------|------|------|------|------|------|
| 120 образов, % распознавания | 100 | 100 | 99,1 | 98,3 | 96,6 | 96,6 | 95,0 | 95,8 | 88,3 |
| Два из трех, % распознавания | 100 | 100 | 100  | 100  | 100  | 100  | 100  | 97,5 | 95,0 |

результата распознавания по двум правильно узванным образам из трех приведена на рис. 5.5.13 (правая сторона).

Аналогичные вышеприведенным оценки были получены и для других значений числа классов  $K$  и других параметров систем распознавания.

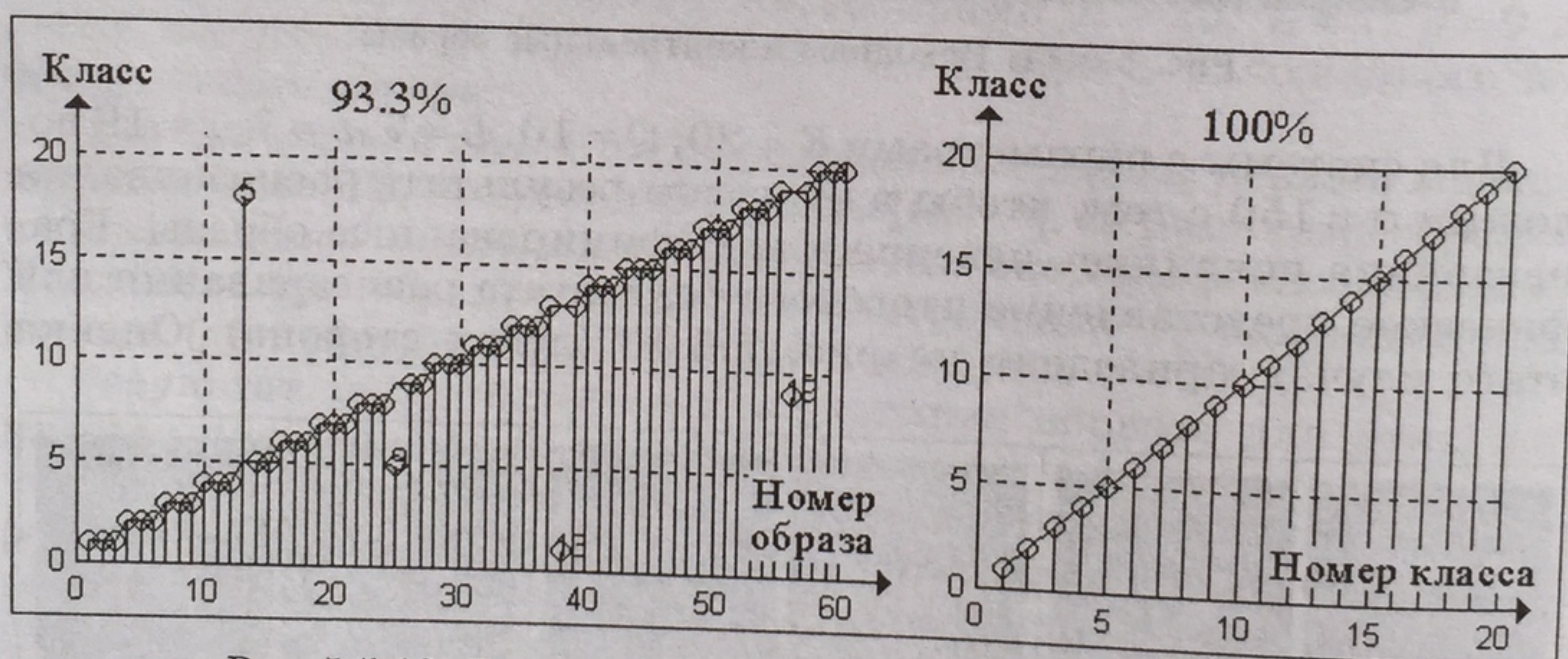


Рис. 5.5.13. Два способа оценки результата распознавания

Таким образом, граница для  $\sigma$ , при которой возможно правильное распознавание зашумленных контрольных образов, находится в пределах до 100.

При этом следует отметить, что, во-первых, в практике распознавания не используют таких высоких значений  $\sigma$  наложенного шума, а во-вторых, в данном случае не была выполнена предварительная фильтрация распознаваемых образов.



### 5.5.3. РАСПОЗНАВАНИЕ ПРИ МАСШТАБИРОВАНИИ ОБРАЗОВ

На практике при распознавании человека по лицу часто распознаваемый образ (исходный образ лица) меньше по размеру, чем требуется. Это происходит, например, при фотографировании некоторой сцены на отдаленном расстоянии и выделении конкретного образа лица из сцены, выделении образа лица из различных фотодокументов и т. д. [Sch99].

В исследуемом нами методе исходными признаками являются яркости всех  $MN$  пикселей распознаваемого образа, а редукция пространства признаков реализуется в соответствии с матрицами  $F_{p \times M}^{(R)}$  и  $F_{N \times p}^{(C)}$ , ориентированными на заданный размер изображения, содержащего  $M$  строк и  $N$  столбцов. Очевидно, что в данном случае необходимо перемасштабирование исходного (уменьшенного) изображения до заданного размера. Для приведения исходного изображения к заданному размеру используется двумерная интерполяция, например, бикубическая, билинейная или так называемая «ступенчатая». Последний вид интерполяции сводится, по существу, к соответствующему повторению каждого пикселя исходного изображения  $M/t$  раз по строкам и  $N/n$  по столбцам, где  $t$  и  $n$  — число строк и столбцов исходного изображения.

Каждый вид интерполяции вносит определенные искажения в результирующий образ. В лучшем случае можно получить как бы «расфокусированное» изображение, в худшем — «грубо квантованное» (рис. 5.5.14). На этом рисунке представлен контрольный образ ВВ, размер которого —  $112 \times 92$  пикселя. Размер этого образа был уменьшен в 6 раз, а затем восстановлен различными видами интерполяции, для чего были использованы процедуры `imresize` пакета MATLAB.



Рис. 5.5.14. Искажения, возникающие при перемасштабировании образов



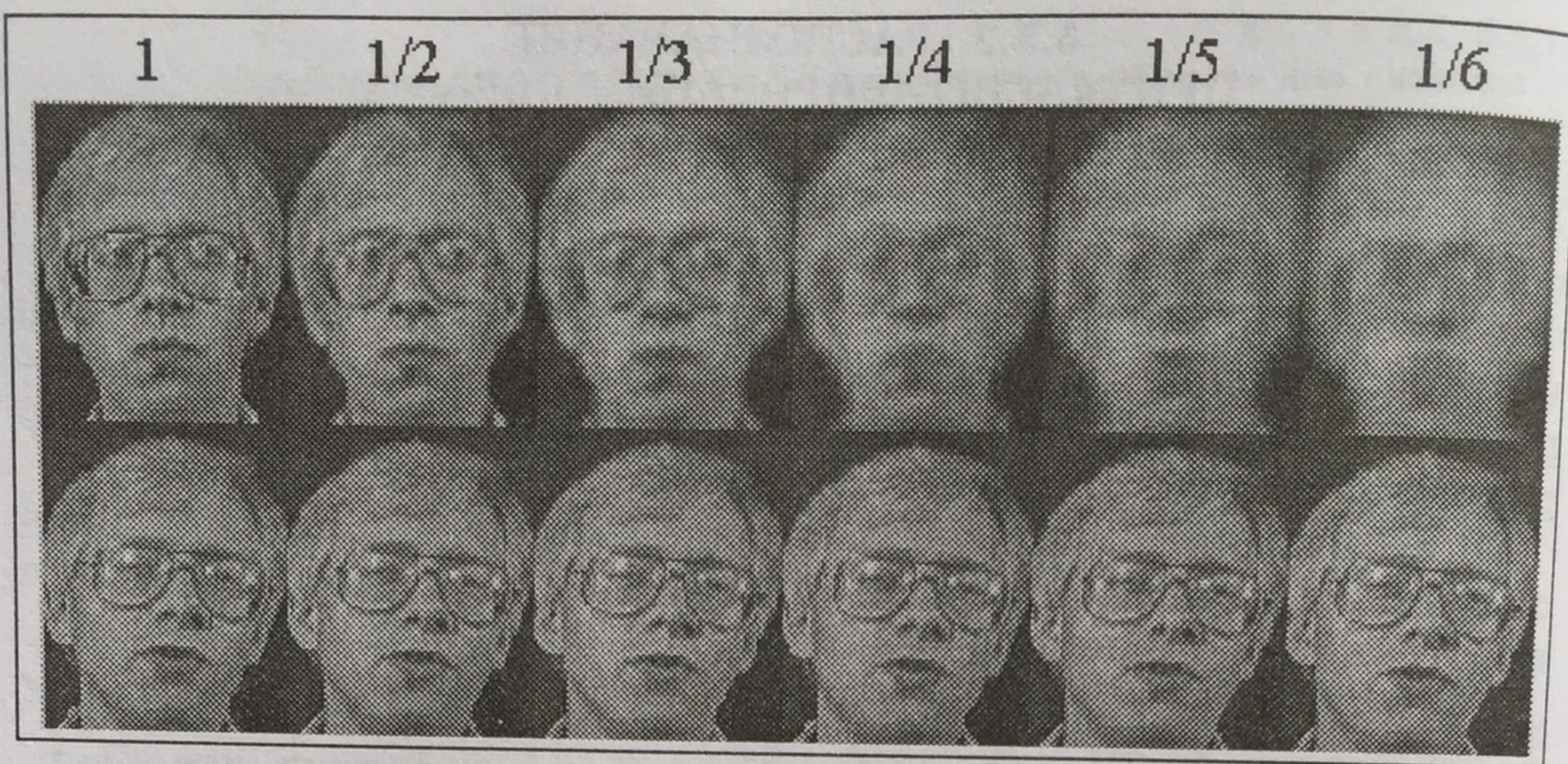


Рис. 5.5.15. Некоторые результаты распознавания контрольных образов при изменении их масштаба и возникающем при этом эффекте их расфокусировки

Используем такое двойное перемасштабирование для имитации реальной ситуации, возникающей в задачах распознавания, и оценим качество распознавания получаемых при этом образов.

**Распознавание расфокусированных образов.** Изменим размеры контрольных образов ВБ соответственно в 2, 3, 4, 5, 6 и 7 раз, а затем восстановим их первоначальные размеры. При использовании бикубической и билинейной интерполяции достигнем эффекта «расфокусировки».

Результат одного из вариантов распознавания для такого случая показан на рис. 5.5.15, где в верхней строке указаны коэффициенты изменения масштаба, во второй строке — образы, предъявленные для распознавания, а в третьей — наиболее близкие образы, выбранные из ВБ в процессе распознавания расфокусированных контрольных образов.

Итоговые результаты распознавания 3 контрольных образов в каждом из 40 классов (для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$ ) приведены во второй строке табл. 5.5.3 и 5.5.4, в строке 3 указана оценка качества распознавания по мажоритарному принципу.

**Распознавание грубо квантованных образов.** Эффекта грубого квантования контрольного образа можно достичь при исполь-

Т а б л и ц а 5.5.3  
Итоговые результаты распознавания 120 контрольных образов для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$  и бикубической интерполяции

| Масштаб                          | 1   | 1/2 | 1/3 | 1/4   | 1/5   | 1/6  | 1/7  |
|----------------------------------|-----|-----|-----|-------|-------|------|------|
| Распознавание 120 образов, %     | 100 | 100 | 100 | 99,16 | 98,33 | 92,5 | 96,6 |
| Распознавание по двум из трех, % | 100 | 100 | 100 | 100   | 100   | 95,0 | 100  |



Т а б л и ц а 5.5.4

Итоговые результаты распознавания 120 контрольных образов для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$  и билинейной интерполяции

| Масштаб                          | 1   | 1/2 | 1/3 | 1/4 | 1/5  | 1/6  | 1/7  |
|----------------------------------|-----|-----|-----|-----|------|------|------|
| Распознавание 120 образов, %     | 100 | 100 | 100 | 100 | 97,5 | 92,5 | 91,6 |
| Распознавание по двум из трех, % | 100 | 100 | 100 | 100 | 97,5 | 95,0 | 92,5 |

зовании ступенчатой интерполяции. Здесь также изменим размеры контрольных образов ВБ соответственно в 2, 3, 4, 5, 6 и 7 раз, а затем восстановим их до первоначальных.

Результат одного из вариантов распознавания для такого случая приведен на рис. 5.5.16, где в верхней строке указан коэффициент изменения масштаба, во второй строке — образы, предъявленные для распознавания, а в третьей — наиболее близкие образы, выбранные из ВБ в процессе распознавания расфокусированных контрольных образов.

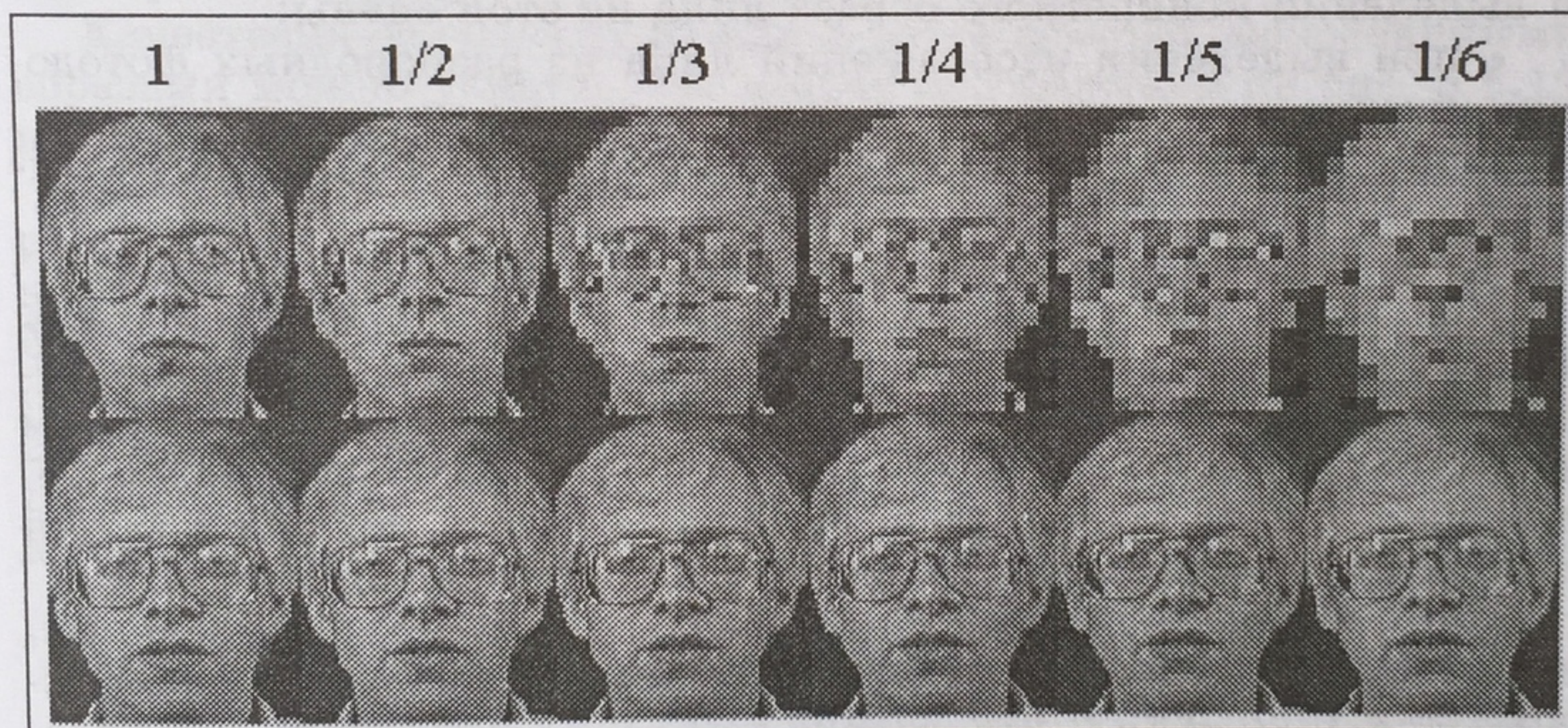


Рис. 5.5.16. Некоторые результаты распознавания контрольных образов при изменении их масштаба и возникающем при этом эффекте их грубого квантования

Т а б л и ц а 5.5.5

Итоговые результаты распознавания 120 контрольных образов для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$  и ступенчатой интерполяции

| Масштаб                          | 1   | 1/2 | 1/3 | 1/4   | 1/5   | 1/6  | 1/7   |
|----------------------------------|-----|-----|-----|-------|-------|------|-------|
| Распознавание 120 образов, %     | 100 | 100 | 100 | 99,16 | 98,33 | 97,5 | 98,33 |
| Распознавание по двум из трех, % | 100 | 100 | 100 | 100   | 97,5  | 97,5 | 100   |



Итоговые результаты распознавания 3 контрольных образов в каждом из 40 классов (для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$ ) представлены во второй строке табл. 5.5.5, в строке 3 — оценка качества распознавания по мажоритарному принципу.

Как показывают полученные результаты, вид интерполяции, используемый при восстановлении заданного размера контрольного образа, незначительно влияет на качество распознавания. Из этих же результатов следует, что 100 %-ное распознавание контрольных образов возможно при изменении их масштабов не более чем в 3 – 4 раза. При этом размеры распознаваемых образов не должны быть ниже  $28 \times 23$  пикселя.

#### 5.5.4. РАСПОЗНАВАНИЕ ПРИ МАСШТАБИРОВАНИИ И НАЛОЖЕНИИ ШУМА

Предназначенный для распознавания исходный образ может быть получен из различных источников и в том числе:

- при наблюдении некоторой сцены с отдаленного расстояния и выделении конкретного образа лица из этой сцены;
- при выделении изображений лица из разнородных фотодокументов;
- при выборе образов из компьютерных баз данных, структурированных различными способами;
- непосредственно с цифровой или аналоговой видеокамеры.

При выборе исходного образа из различных источников, его трансмиссии и предварительной обработки (включающей, например, декомпрессию, когда данные в некоторой базе находятся в сжатом виде, процесс оптического или электронного сканирования, дискретизацию и т. д.) размер распознаваемого образа может оказаться меньше, чем требуется и, кроме того, находиться под шумом.

Предвидеть все варианты масштабных и шумовых изменений, возникающих в распознаваемых образах, практически невозможно. Однако положим, что механизмы таких изменений описываются схемами, представленными в табл. 5.5.6, где знаки «+» обозначают наложение шума, а стрелки «↑» или «↓» — перемасштабирование размера исходного образа в сторону увеличения или уменьшения соответственно.

Если говорить о степени искажений, то при одних и тех же значениях  $\sigma$  и масштабном коэффициенте  $\alpha$  при второй схеме из табл. 5.5.6 создаются условия для появления меньших искажений, чем при третьей, но больших, чем при первой. Сказанное касается также вариантов 5, 4 и 6. Отметим, что на практике распознавания образов вторая и пятая схемы искажений встречаются наиболее часто.

С учетом сказанного представим результаты работы системы распознавания для случая, когда на исходный образ сначала нало-



Т а б л и ц а 5.5.6

Механизмы изменения контрольных образов

| № варианта | Размер исходного образа | Шум | Перемасштабирование | Шум |
|------------|-------------------------|-----|---------------------|-----|
| 1          | Меньше требуемого       |     | ↑                   | +   |
| 2          | Меньше требуемого       | +   | ↑                   |     |
| 3          | Меньше требуемого       | +   | ↑                   | +   |
| 4          | Больше требуемого       |     | ↓                   | +   |
| 5          | Больше требуемого       | +   | ↓                   |     |
| 6          | Больше требуемого       | +   | ↓                   | +   |

жен нормальный белый шум, а затем размер этого образа перемасштабируется до требуемого. Таким образом, рассмотрим вторую схему (из табл. 5.5.6) искажений распознаваемых образов, а в качестве исходных используем контрольные образы, размеры которых уменьшены в 2, 3 и 4 раза. При этом масштабный коэффициент  $\alpha$  увеличения размера исходных образов составит 2, 3 и 4.

Качественную границу между верно и неверно распознанными образами можно увидеть, сравнивая изображения на рис. 5.5.17 и 5.5.18 – 5.5.20.

На рис. 5.5.17 в правом нижнем углу выделены три контрольных образа, а на рис. 5.5.18 – 5.5.20 те же образы, но искаженные в соответствии со схемой 2 табл. 5.5.6. Образы на рис. 5.5.18 – 5.5.20 отличаются только видом интерполяции, используемой при перемасштабировании. На каждом из рис. 5.5.18 – 5.5.20 с левой стороны представлены образы (для значений  $\alpha = 4$  и  $\sigma = 20$ ), верно распознанные системой с параметрами  $K = 20$ ;  $L = 7$ ;  $p = 9$  и  $s = 19$ , а с правой (для значений  $\alpha = 4$  и  $\sigma = 40$ ) — неверно распознанные.



Рис. 5.5.17. Контрольные образы 20-го класса ВВ



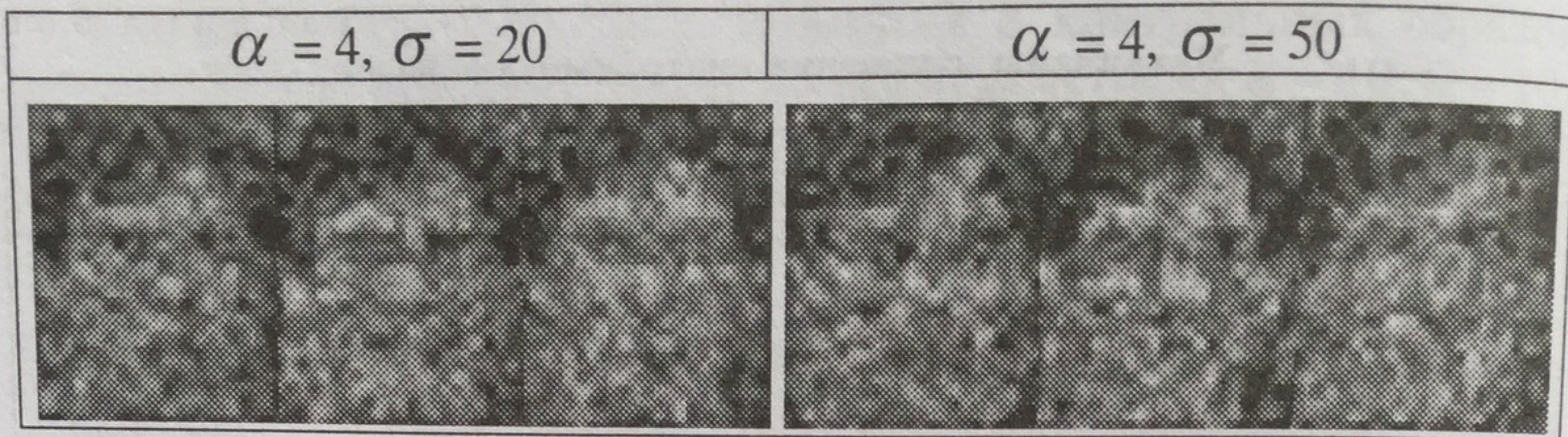


Рис. 5.5.18. Искажения контрольных образов (бикубическая интерполяция)

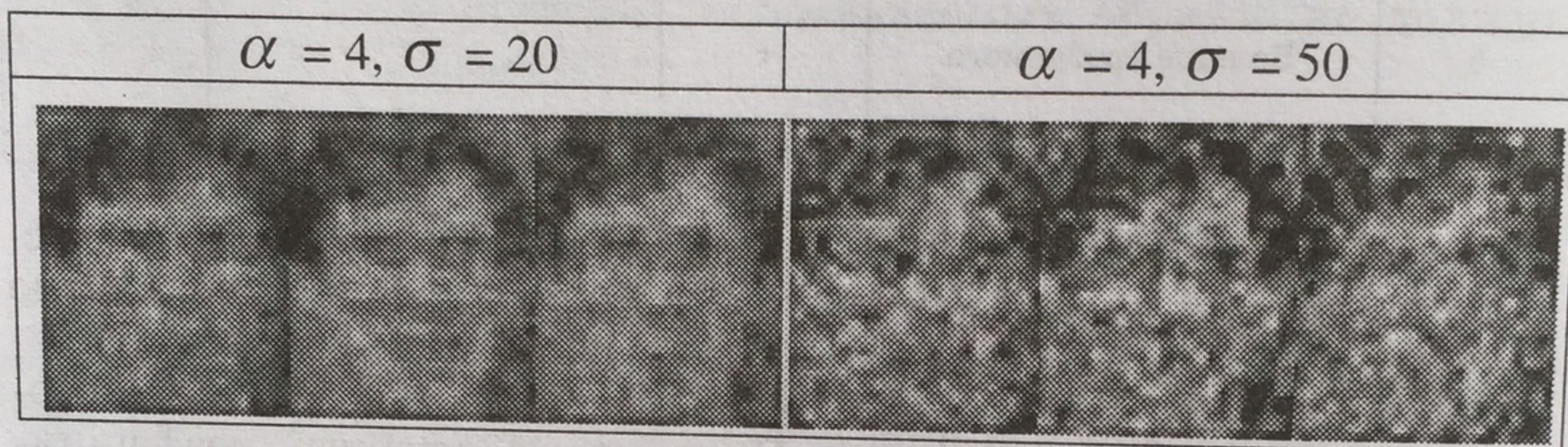


Рис. 5.5.19. Искажения контрольных образов (билинейная интерполяция)

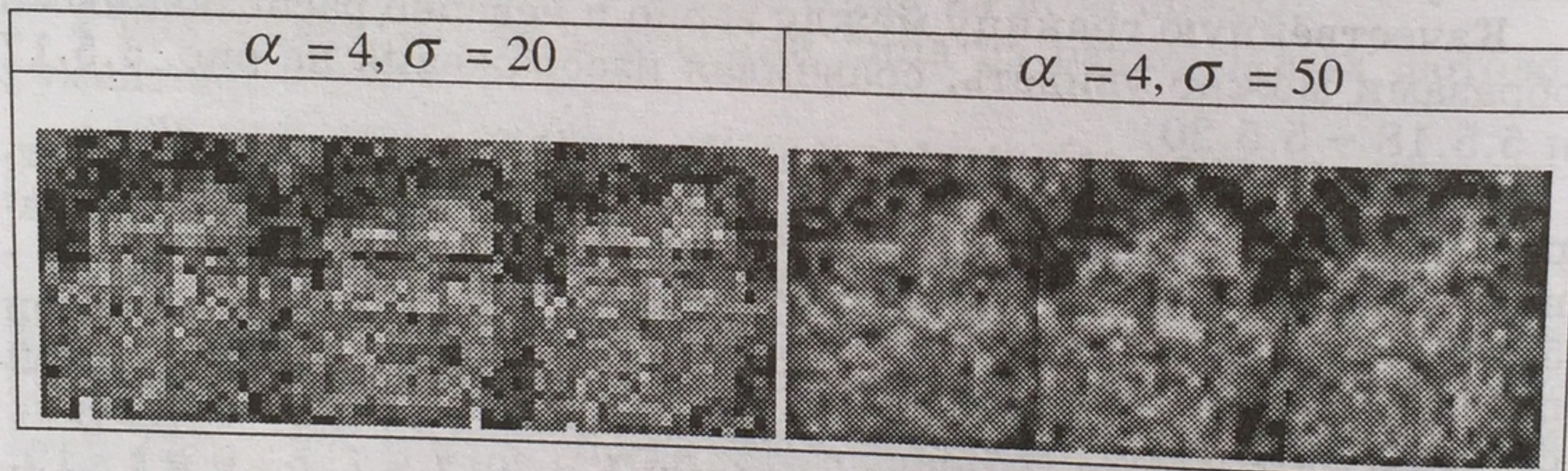


Рис. 5.5.20. Искажения контрольных образов (ступенчатая интерполяция)

В проведенном эксперименте целью было определение границ для значений  $\sigma$  и масштабного коэффициента  $\alpha$ , при которых выполняется верное распознавание искаженных образов.

Итоговые результаты распознавания 3 контрольных образов в каждом из 20 классов для системы с параметрами  $K = 20, Q = 10, L = 7, p = 7, s = 19$  и бикубической интерполяции представлены в табл. 5.5.7 (третья строка).

В четвертой строке табл. 5.5.7 приведены итоговые результаты распознавания по мажоритарному принципу. Как видим, в целом верное распознавание возможно при значениях  $\alpha = 3, 4$  и  $\sigma = 20$ .

Итоговые результаты распознавания 3 контрольных образов в каждом из 20 классов для системы с параметрами  $K = 20, Q = 10, L = 7, p = 7, s = 19$  и билинейном способе интерполяции в процессе перемасштабирования образов приведены в табл. 5.5.8 (третья и четвертая строки).



Т а б л и ц а 5.5.7

Итоговые результаты распознавания 60 контрольных образов для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  и бикубической интерполяции

| Масштаб                                  | 1/2 |     |     |     | 1/3 |     |     |    | 1/4 |      |      |      |
|--|-----|-----|-----|-----|-----|-----|-----|----|-----|------|------|------|
| $\sigma$                                 | 20  | 30  | 40  | 50  | 20  | 30  | 40  | 50 | 20  | 30   | 40   | 50   |
| Распознавание 60 образов, %              | 100 | 100 | 100 | 100 | 100 | 100 | 98  | 90 | 100 | 96,6 | 91,6 | 73,3 |
| Распознавание по двум образам из трех, % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 90 | 100 | 100  | 95,0 | 80   |

Т а б л и ц а 5.5.8

Итоговые результаты распознавания 60 контрольных образов для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  и билинейной интерполяции

| Масштаб                                  | 1/2 |     |     |     | 1/3 |     |      |      | 1/4 |      |      |      |
|--|-----|-----|-----|-----|-----|-----|------|------|-----|------|------|------|
| $\sigma$                                 | 20  | 30  | 40  | 50  | 20  | 30  | 40   | 50   | 20  | 30   | 40   | 50   |
| Распознавание 60 образов, %              | 100 | 100 | 100 | 100 | 100 | 100 | 96,6 | 93,3 | 100 | 96,6 | 91,6 | 81,6 |
| Распознавание по двум образам из трех, % | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100 | 100  | 100  | 90   |

Т а б л и ц а 5.5.9

Итоговые результаты распознавания 60 контрольных образов для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  и ступенчатой интерполяции

| Масштаб                                  | 1/2 |     |     |     | 1/3 |     |     |     | 1/4 |      |      |      |
|--|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|
| $\sigma$                                 | 20  | 30  | 40  | 50  | 20  | 30  | 40  | 50  | 20  | 30   | 40   | 50   |
| Распознавание 60 образов, %              | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 95  | 100 | 96,6 | 93,3 | 91,6 |
| Распознавание по двум образам из трех, % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 95,0 | 100  |

Как видим, и здесь в целом верное распознавание возможно при значениях  $\alpha = 3$  и  $\sigma = 30$ , а также при значениях  $\alpha = 4$  и  $\sigma = 20$ .

Итоговые результаты распознавания 3 контрольных образов в каждом из 20 классов для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  и ступенчатым способом интерполяции в процессе перемасштабирования образов приведены в табл. 5.5.9 (третья и четвертая строки).

Последний эксперимент показывает, что верное распознавание искаженных масштабированием и шумом образов возможно при значениях  $\alpha = 3$  и  $\sigma = 40$ , а также при значениях  $\alpha = 4$  и  $\sigma = 20$ . Однако граница для параметра  $\sigma$  скорее всего не должна превышать величины 20.





Рис. 5.5.21. Исходные и контрольные образы 40-го класса ВВ

О подобных границах параметров  $\alpha$  и  $\sigma$  могут свидетельствовать и результаты распознавания контрольных образов в базе данных, содержащей 40 классов образов. На рис. 5.5.21 представлены все 10 образов 40-го класса, а контрольные образы выделены в нижнем правом углу.

Изображения, предназначенные для распознавания и искаженные в соответствии со схемой 2 табл. 5.5.6, представлены на рис. 5.5.22.

Итоговые результаты распознавания 3 контрольных образов в каждом из 40 классов для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 39$  и ступенчатым способом интерполяции в процессе перемасштабирования образов представлены в табл. 5.5.10 (третья и четвертая строки). Из этих результатов видно, что в целом верное распознавание контрольных образов возможно при  $\alpha = 2, 3, 4$  и  $\sigma = 20$ .

Все выполненные эксперименты по распознаванию контрольных образов при их масштабировании и наложении на них шума показывают, что устойчивость процесса распознавания высокая.

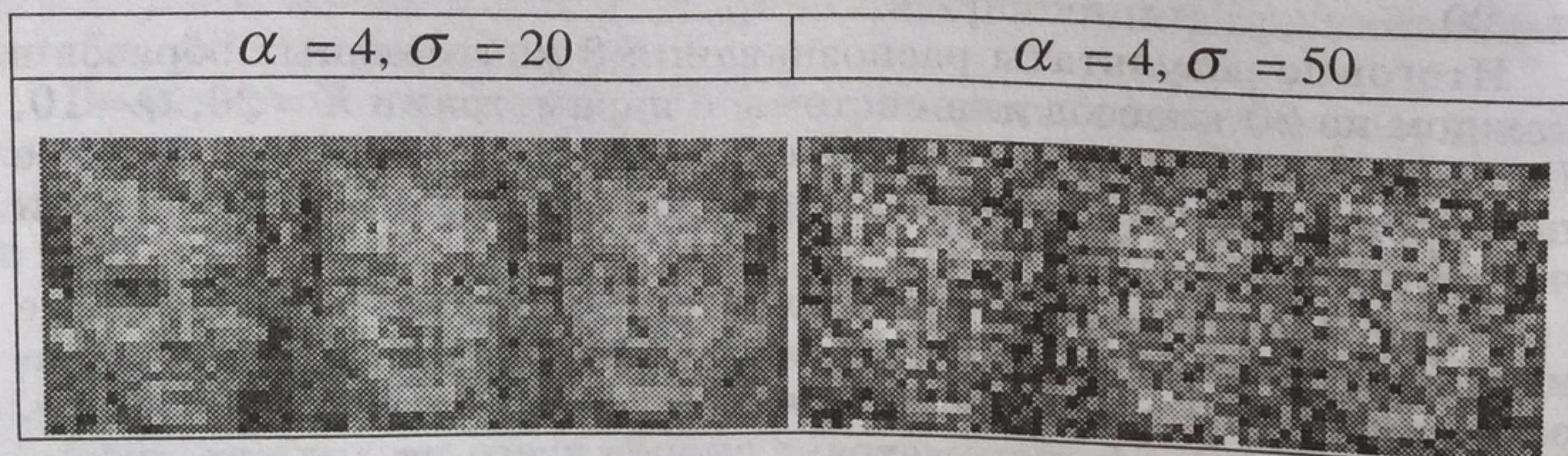


Рис. 5.5.22. Искажения контрольных образов (ступенчатая интерполяция)



Т а б л и ц а 5.5.10

Итоговые результаты распознавания 120 контрольных образов для системы с параметрами  $K = 40$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 12$ ,  $s = 19$  и ступенчатой интерполяцией

| Масштаб                                  | 1/2 |     |     |     | 1/3 |      |      |      | 1/4 |      |      |      |
|--|-----|-----|-----|-----|-----|------|------|------|-----|------|------|------|
| $\sigma$                                 | 20  | 30  | 40  | 50  | 20  | 30   | 40   | 50   | 20  | 30   | 40   | 50   |
| Распознавание 120 образов, %             | 100 | 100 | 98  | 97  | 100 | 96,6 | 94,1 | 85,8 | 100 | 93,3 | 86,6 | 74,1 |
| Распознавание по двум образам из трех, % | 100 | 100 | 100 | 100 | 100 | 100  | 97,5 | 90   | 100 | 97,5 | 92,0 | 82   |

Достижение 100 %-ного уровня распознавания контрольных образов возможно при значительных изменениях параметров  $\alpha$  и  $\sigma$  и достаточных для решения практических задач. Таким образом, можно подтвердить качество метода распознавания, рассмотренного в настоящей главе и реализованного на основе работ [Tsa98, Swe96,99].

#### 5.5.5. РАСПОЗНАВАНИЕ ОБРАЗОВ, НЕ ПРИНАДЛЕЖАЩИХ СОБСТВЕННОЙ БАЗЕ

А теперь зададимся таким вопросом: что произойдет, если для распознавания будут предъявлены образы, не входящие в исходную базу данных?

Ясно, что на каждый предъявленный к распознаванию образ система подберет наиболее близкий ему из образов, имеющих в базе. Однако возникают при этом и другие вопросы: какие при этом получатся результаты и можно ли по ним определить, что предъявленный образ не принадлежит данной базе? Что будет определять близость образов и что будет критерием их различия?

Для получения ответов на поставленные вопросы выполним следующий эксперимент. Предъявим системе с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  60 собственных контрольных образов и 87 новых, из которых 60 принадлежат классам 21 – 40 базы данных [Orl\*\*], а 27 — базе, использованной в [Kuc99] и частично представленной на рис. 5.5.8. В дальнейшем все эти новые образы также будем называть контрольными.

За качество распознавания примем меру подобия контрольного образа тому классу ВБ, который будет выбран системой, и эту меру оценим с помощью метрик  $L_0$  и  $L_2$ . Иными словами, будем вычислять расстояния между признаками контрольного образа и средними признаками выбранного системой класса (все оценки выполним в редуцированном пространстве признаков). Для метрики  $L_0$  примем значение порога  $\delta = 30$ , что составляет не более 30 % максимальных значений признаков. Далее, поскольку из



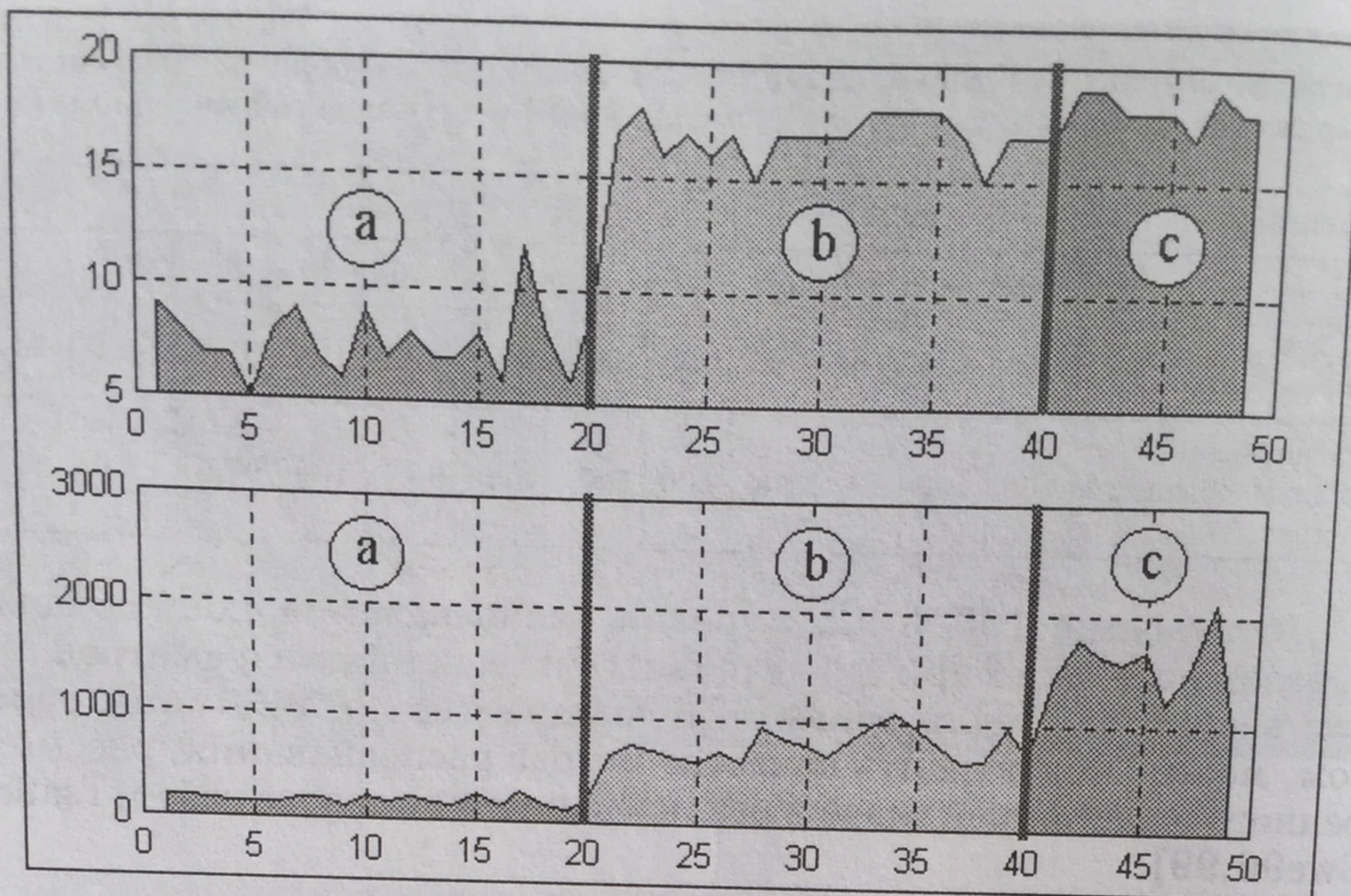


Рис. 5.5.23. Результаты оценки расстояний для «своих» и «чужих» образов

каждого класса системе будут предъявлены по три контрольных образа, результаты оценки расстояний усредняются по трем индивидуальным оценкам.

В результате выполненного эксперимента (или «распознавания»?) получены по 49 оценок расстояний для каждой из метрик. Эти оценки представлены на рис. 5.5.23. Верхний график на рисунке характеризует оценки расстояний по метрике  $L_0$ , а нижний — по метрике  $L_2$ . Буквами а, b и с обозначены соответственно зоны:

а) — оценки расстояний для контрольных образов, принадлежащих ВБ;

б) — оценки расстояний для контрольных образов базы [Orl\*\*] — классы с 21-го по 40-й;

с) — оценки расстояний для контрольных образов, принадлежащих базе [Kuc99] — классы с 41-го по 49-й.

Представленные результаты убедительно показывают следующее:

- наименьшие расстояния получены для контрольных образов, принадлежащих собственной ВБ (зона а графиков);

- оцениваемые расстояния хорошо разделяют контрольные образы на «свои» и «чужие», т. е. образы, принадлежащие собственной ВБ и не принадлежащие ей;

- оцениваемые расстояния (по крайней мере, в метрике  $L_2$ ) также хорошо разделяют зоны b и c графиков, т. е. контрольные образы, полученные из разнородных источников.

Метрика  $L_0$  не выявила различий в контрольных образах, полученных из разнородных источников, по двум причинам: поскольку, во-первых, она не приспособлена для оценки расстояний в пространстве редуцированных признаков, что отмечено в [Sim00], а во-





Рис. 5.5.24. Результаты подбора к контрольному образу образа из ВБ

вторых, эксперимент проводился для системы с параметрами  $K = 20$ ,  $Q = 10$ ,  $L = 7$ ,  $p = 7$ ,  $s = 19$  и для оценки расстояний использовалось всего 19 компонент. Хотя и для этой метрики разделение контрольных образов на «свои» и «чужие» тоже заметно (в зоне а расстояния существенно меньше, чем в зонах (b + c).

В целом же для целей верификации распознанного образа — «свой» или «чужой» — можно, очевидно, использовать обе метрики, хотя в данном случае предпочтение следует отдать метрике  $L_2$ .

Оценим теперь результаты выполненного эксперимента с качественной стороны, для чего сравним визуально между собой контрольный образ, не принадлежащий ВБ, и образ, выбранный системой. На рис. 5.5.24 приведены три пары образов, из которых каждый левый образ в парах контрольный, а правый — образ, выбранный системой. В верхней строке над каждым образом указаны номера классов, которым они принадлежат (в нумерации базы [Orl\*\*]).

Визуальный анализ приведенных пар образов подтверждает — система распознавания действительно выбирает из ВБ образ, наиболее близкий контрольному: в каждой паре образов можно найти определенную меру сходства.

В целом же при отсутствии в ВБ «очень похожего» образа система подберет образ, в котором сходство с контрольным образом будет определяться: одинаковым поворотом и/или наклоном головы, общими элементами на лице (брови, глаза, очки, борода, вид прически, цвет волос и т. д.), выражением лица, фоном, на котором размещено лицо (черный, серый, белый), размещением собственно лица в поле всего изображения (циклический сдвиг) и т. д.



## 6. СИСТЕМЫ РАСПОЗНАВАНИЯ «ON-LINE»

### 6.1. СТРУКТУРА СИСТЕМЫ РАСПОЗНАВАНИЯ

Ниже будут представлены системы, работающие в отличие от рассмотренных в гл. 5 с цифровой камерой в режиме «on-line». При этом решается ряд дополнительных проблем, которые связаны с необходимостью предварительной обработки изображений с камеры и включают:

- локализацию требуемого лица из окружающего его фона или отделение его от других лиц;
- выделение наиболее информативной части лица из его общего вида в соответствии с выбранной частью на эталоне (например, 2/3 нижней части лица, включая брови, глаза, нос, рот);
- согласование размеров лиц (для достижения одинакового масштаба с эталоном);
- согласование пункта наблюдения распознаваемого лица с выбранным для эталона (элиминация больших углов поворота головы в момент получения образа лица) и т. д.
- элиминацию влияния мимики, прически или других элементов (серьжки и кольца, очки и т. д.) распознаваемого образа;
- элиминацию способа освещения данного образа или согласование с освещением для выбранного эталона.

Отмеченные проблемы становятся особенно актуальными в тех случаях, когда распознавание выполняется по небольшому числу эталонов в каждом классе образов, что обычно и имеет место на практике. И хотя большая часть известных систем распознавания «on-line» основана на использовании баз данных, в которых имеются различные варианты образов лица в каждом классе, — невозможно предусмотреть все случаи, встречающиеся в жизни (т. е. все варианты масштабов исходных изображений, поворотов головы — пунктов наблюдения, освещения, фона и т. д.). Практически каждый раз мы имеем дело с распознаванием нового образа, который как бы не совсем точно соответствует собственной базе данных (см. конец гл. 5).

Процесс распознавания в системе «on-line» можно представить следующим образом. Пусть несколько разных изображений лиц (образов) заданного класса (или классов) находятся в базе данных. Каждый такой образ можно представить как эталон. В процессе реализации задачи распознавания новый образ получаем непосредственно с камеры. Необходимо проверить его принадлежность базе данных: либо посредством вычисления некоторой меры подобия между новым образом и каждым эталоном, либо посредством вычисления меры подобия между некоторой характеристикой нового образа и совместной характеристикой в классе образов. Новый образ будет принадлежать тому классу, мера подобия с которым будет наивысшей и если эта мера превышает некоторый порог идентификации.





Рис. 6.1.1. Процесс распознавания нового образа в системе «Visitor Identification»

Трудности реализации подобных процессов распознавания покажем на примере системы «Visitor Identification», реализующей идеи работы [Sim00].

Системы класса «Visitor Identification» решают задачу идентификации людей по изображению лица (например, идентификация гостя, входящего в дом или бюро, верификация посетителя страницы в Интернете, распознавание постоянного покупателя в магазине и т. д.). Главная особенность таких систем (при решении самой задачи распознавания человека по лицу) — простота реализации задач и работа в режиме «on-line».

Основные операции обработки образов в системе «Visitor Identification» представлены на рис. 6.1.1. Исходная база данных содержит 10 классов образов — по одному образу в каждом классе. Все образы базы представлены в нижней части рисунка. Размер изображения, полученного с камеры,  $240 \times 320$  пикселей. Интересующий нас образ лица автоматически ищется в поле выделенной рамки (левый верхний кадр) с помощью вспомогательного эталона, содержащего произвольное лицо на белом фоне. В следующем кадре видно завершение поиска интересующего нас образа. По окончании процесса выделения лица из исходного изображения получаем образ для распознавания размером  $112 \times 96$



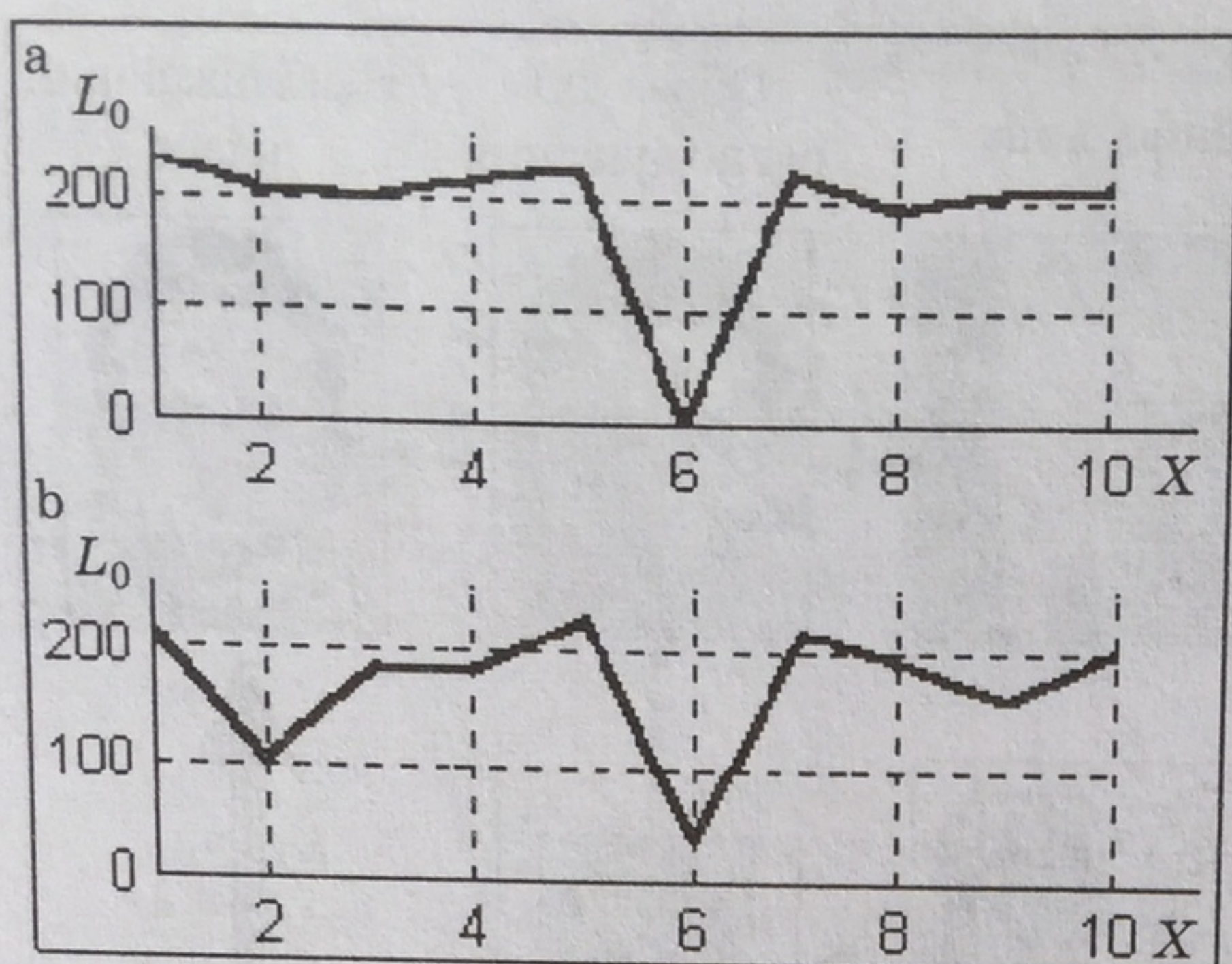


Рис. 6.1.2. Функции превышения порога: а — функция превышения порога для образа, взятого из базы данных; б — функция превышения порога для нового образа, полученного с камеры

пикселей (третий кадр сверху).

Идея подхода [Sim00] состоит в том, что образы исходной базы данных и распознаваемый образ масштабируются к размеру  $16 \times 16$  пикселей

(результат такого масштабирования представлен на рис. 6.1.1 в блоке за стрелкой «Модификация»).

При этом элиминируются такие нежелательные факторы, как масштаб (в пределах  $\pm 10\%$ ), циклический сдвиг собственно лица в поле рамки целого образа (в пределах  $1 - 2\%$  от размера рамки), повороты головы (в пределах  $15\%$ ). Далее распознаваемый образ (размера  $16 \times 16$  пикселей) сравнивается с каждым перемасштабированным образом базы.

При этом используется метрика:

$$L_0 \equiv \sum 1 \quad \text{при условии } |x_{ij} - y_{ij}| > \delta, \quad (6.1.1)$$

по которому каждый пиксель  $x_{ij}$  одного образа сравнивается с соответствующим пикселем  $y_{ij}$  другого образа, и, если модуль их разности превышает некоторый заданный порог  $\delta$ , то параметр меры превышения порога («меры неподобия») увеличивается на единицу.

В результате сравнений с образами всей базы получается некоторый ряд, элементами которого будут параметры мер превышения порога. Решение о принадлежности распознаваемого образа к определенному классу принимается по минимуму значения этой меры.

Функция превышения порога для образа 6-го класса, взятого непосредственно из базы данных, приведена на рис. 6.1.2 а. На оси  $X$  отмечены номера классов, а по оси  $Y$  — показатель меры превышения порога. Видно, что *минимум* приходится на шестой класс. Результат сравнения образа, полученного с камеры (также 6-й класс), со всеми 10 образами базы показан в форме функции превышения порога на рис. 6.1.2 б. Здесь также видно, что *минимум* приходится на шестой класс, хотя форма всей функции уже не такая «идеальная», как выше.

Если результат, полученный при сравнении, нормировать на общее число пикселей сравниваемых изображений (в данном случае на  $16 \times 16 = 256$ ) и отнять его от единицы, то получим меру подобия, выраженную в диапазоне от 0 до 1. Ясно, что чем ближе



эта мера к единице, тем выше вероятность правильного распознавания. На рис. 6.1.3 приведена нормированная мера для рассмотренного выше примера. Из него также видно, что **максимум** приходится на шестой класс.

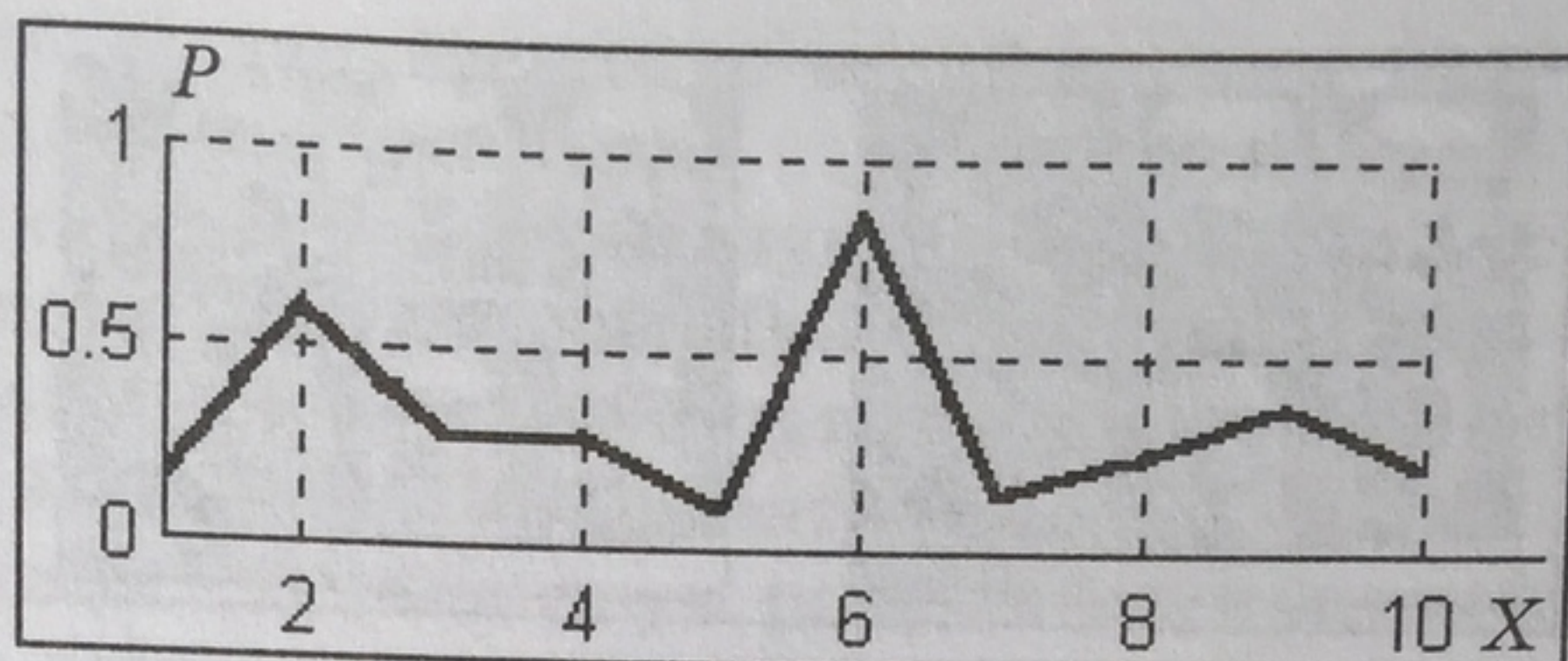


Рис. 6.1.3. Нормированная мера показателя подобия  $P = 1 - (\Sigma 1)/256$  для нового образа шестого класса, полученного с камеры

А теперь обратим внимание на качество распознавания: нормированная мера показателя подобия во втором случае была чуть ниже 1, что можно классифицировать как «неплохое распознавание». Однако здесь видны два «претендента» на признание принадлежности распознаваемого образа — к классу 2 (с мерой подобия выше 0,5) и к классу 9 (с мерой подобия несколько ниже 0,5). Если внимательно посмотреть на образы исходной базы (см. рис. 6.1.1), то можно найти много общего между образами 6-го и 2-го, а также 6-го и 9-го классов.

Предъявим системе новые образы, причем такие, которые или не входят в базу данных или не относятся к фотографиям лица человека. Эти новые образы и выбранные системой претенденты приведены на рис. 6.1.4 и 6.1.5.

Анализируя идею работы [Sim00] и представленные результаты, можно отметить:

- система реализует очень простой алгоритм распознавания и, следовательно, не требует сложных вычислений и больших объемов памяти;
- система идентифицирует образы, относящиеся к базе данных, при их получении практически в тех же условиях, что и образы исходной базы данных, и, кроме того, распознаваемый образ должен быть «актуальным» по отношению к образу в базе данных;
- реализованная система имеет 100 %-ный уровень ложной



Рис. 6.1.4. Образы, не входящие в базу данных, и результат их распознавания





Рис. 6.1.5. Результат распознавания некорректно выбранного образа

идентификации для образов, не относящихся к базе данных, хотя в выбранных системой классах-претендентах можно найти и много общего с предъявленными для распознавания образами: система реагирует на внешнюю форму распознаваемого образа, на овал лица, форму прически и яркость освещения, заданного для распознавания образа.

Для улучшения качества распознавания в [Sim00] предложено проводить распознавание на 3 – 5 образах одного и того же «посетителя» (или хранить в базе данных 3 – 5 вариантов образов для каждого класса), что, однако, гарантирует распознавание в режиме «on-line» не выше 55 %. Таким образом, и в данном случае придется минимизировать или полностью исключить влияние всех мешающих факторов, в том числе постоянно обновлять исходную базу данных «актуальными» образами посетителя. Проблема ложной идентификации образов, не принадлежащих классам исходной базы, возникает не только при использовании рассматриваемой системы, но и других, и решается она либо методами верификации результатов распознавания, либо дополнением исходной базы данных специальными образами — «ловушками» [Кис99а].

Итак, представленная система не дает гарантии идентификации «посетителей». На результат распознавания влияют многочисленные факторы: плохое освещение, размер исходного образа, не соответствующий базовому, дополнительные элементы на лице, неполное соответствие формы и вариантов предъявленных образов образам в исходной базе данных.

Рассмотренный выше пример позволяет сформулировать основные требования к системам распознавания «Visitor Identification», в частности и к системам «on-line» вообще. Рассматриваемый класс систем должен иметь:

- высокое качество распознавания, исключающее ложное принятие решений;
- относительно несложные алгоритмы, обеспечивающие высокое быстродействие решения задачи распознавания;
- простые способы настройки (переналадки) системы под новую базу данных.

Высокое качество распознавания возможно при исчерпывающе полной исходной базе данных и совершенстве используемых методов (см., например, разд. 2), а также при предобработке новых образов (например, специальной нормализации, приводящей их к стандартным образам). Кроме того, высокое качество распознава-



ния может поддерживаться за счет процедур верификации принятых алгоритмом распознавания решений.

Верификация принятых решений реализуется на практике несколькими способами [Lan97, Kuc99a, Kuc00, Pen00, Hon98].

1. Контрольной проверкой/верификацией путем распознавания дополнительного образа или наиболее информативной его части.

2. Использованием в одной системе двух разных алгоритмов распознавания, из которых один всегда основной, а другой дополнительный, исполняющий функции верификатора (например, алгоритмов, основанных на [Sim00] и обычной корреляции).

3. Использованием персональных номеров идентификации (англ. PIN), играющих роль пароля.

4. Комплексованием источников биометрической информации и их обработкой в одной системе распознавания (например, лицо + отпечаток пальца, лицо + голос, лицо + глаз).

5. Использованием (анализом) признаков, относящихся к различным категориям, но принадлежащих одному и тому же распознаваемому образу (например, математические и структурные признаки) и т. д.

С учетом всего отмеченного выше, система распознавания «online» должна иметь разветвленную структуру, один из возможных вариантов которой представлен на рис. 6.1.6. Главными блоками здесь являются:

- 1) цифровая камера — источник исходного изображения;
- 2) блок предварительной обработки изображений с камеры, включающий:
  - блоки выделения собственно лица из целого изображения — исходного образа и блоки выделения деталей лица;
  - блоки нормализации исходного образа и согласования размеров;
  - блоки элиминирования поворотов и наклонов изображения головы и т. д.
- 3) исходная база данных;
- 4) виртуальная база данных;
- 5) блоки выбора признаков;
- 6) блок, реализующий основной алгоритм распознавания;
- 7) блок, реализующий дополнительный алгоритм распознавания;
- 8) блок принятия решений;
- 9) блок ввода PIN или блоки ввода других ключей;
- 10) блок, отображающий результат распознавания;
- 11) блок анализа результатов распознавания.

Исходная база данных комплектуется либо из разнородных источников (фотографии, другие базы и источники данных), либо из изображений, поступающих непосредственно с камеры. Каждое изображение при этом проходит этап предварительной обработки: (выделение изображения лица из исходных изображений; выделение деталей и регионов лица; подготовка различных эталонов —



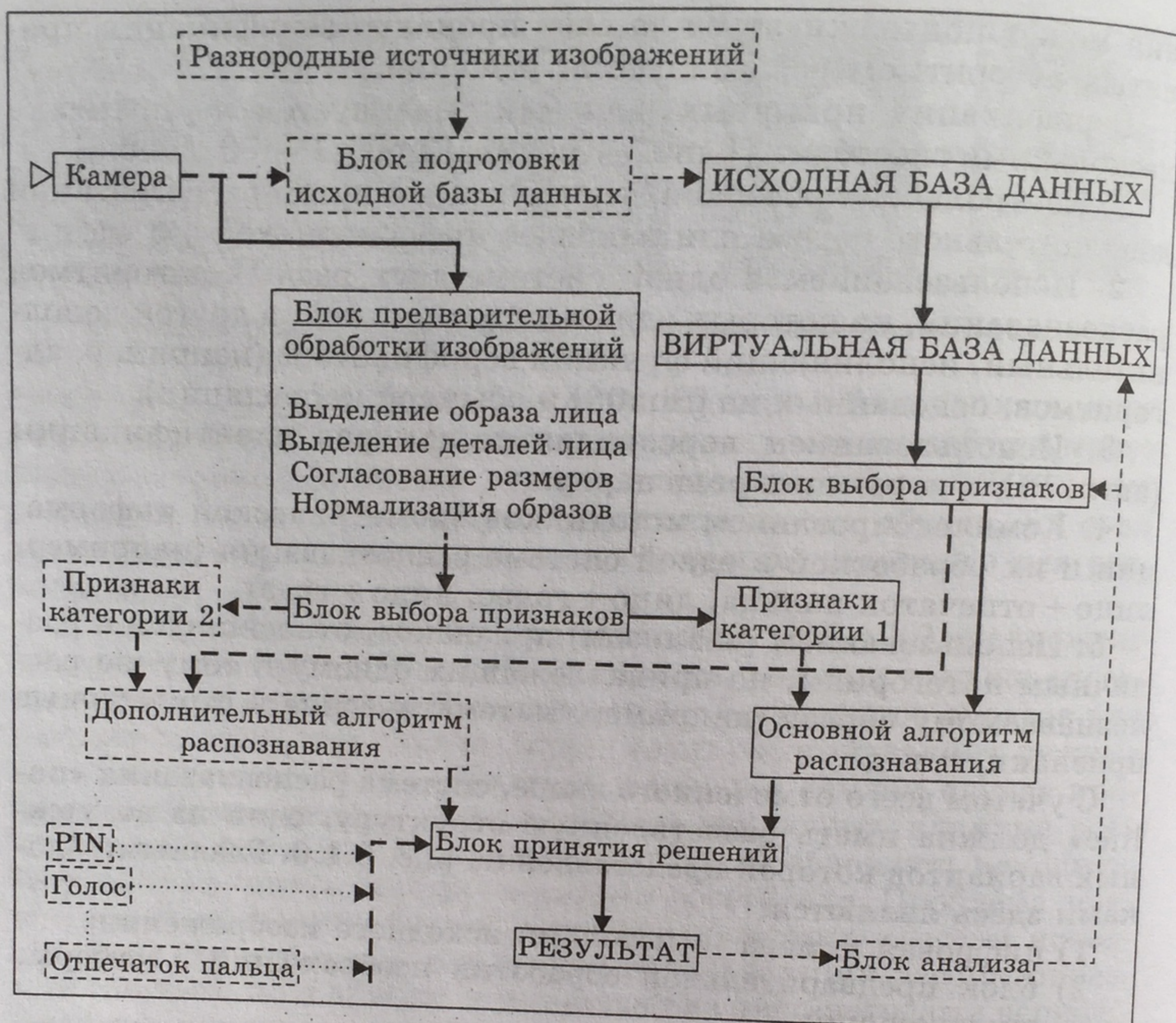


Рис. 6.1.6. Структура системы распознавания «on-line»

средняя часть изображения лица, глаза, нос, рот; согласование размеров выделяемых изображений в соответствии с заданными; нормализация изображений относительно освещения; определение центров глаз, рта, оси симметрии лица и т. д. Сюда же можно отнести и вторичные задачи — определение некоторых параметров собственно лиц, например, расстояния между центрами глаз.

По исходной базе затем выбирается или генерируется виртуальная база (ВБ) данных, в которой определяются образы как для синтеза системы распознавания, так и контрольные образы для ее проверки и оценки качества распознавания непосредственно перед переходом на режим распознавания «on-line». Если в каждом классе исходной базы данных имеется только по одному экземпляру образа, ВБ наполняется подобразами, полученными из имеющихся образов путем их сканирования рамками меньшего размера, выделения отдельных регионов из исходных данных, а также их перемасштабированием (например, для получения пирамиды изображений).

Поскольку в рассматриваемой системе могут быть использованы основной и дополнительный алгоритмы распознавания, выбор



признаков (категория, их параметры, размер признакового пространства) имеет здесь принципиальное значение. Как выбирать признаки и какие, мы уже описывали в разд. 2.3.4 и в гл. 5. Далее опишем выбор признаков применительно ко всем рассматриваемым ниже системам распознавания и с учетом реализуемых в них алгоритмов.

В блоке принятия решений осуществляется вторичная обработка результатов распознавания. Именно здесь принимается решение о повторном (дополнительном) распознавании другого исходного образа в случае неразрешенных альтернатив и включения механизмов дополнительного алгоритма распознавания или ключей контроля результата распознавания. В качестве последних используют отпечатки пальцев, голос, радужную оболочку глаза, руку, PIN, для которых используются специальные алгоритмы и методы распознавания, комплексируемые с основной системой распознавания по лицу.

Блок отображения результата может выдавать результат распознавания в различных аудиовизуальных формах. Главное назначение этого блока — дать сигнал, соответствующий полученному результату распознавания, и активизировать блок анализа и систему в целом на реализацию следующего акта распознавания.

Блок анализа предназначен для накопления информации о результатах распознавания при выбранных параметрах системы и через него могут пополняться базы данных на этапе синтеза системы распознавания.

## 6.2. ВИРТУАЛЬНЫЕ БАЗЫ ДАННЫХ И ВЫБОР ПРИЗНАКОВ

В настоящем разделе сделана попытка обсуждения проблемы неполноты исходных данных и в связи с этим обосновать включение виртуальных баз (ВБ) данных в структуру систем распознавания. Здесь же обсуждаются вопросы о том, какие данные необходимо содержать в ВБ и чем руководствоваться при этом, как и какие выбирать признаки из имеющихся данных для достижения наилучшего результата.

В процессе интеграции данных в исходной базе данных (ИБ), все образы имеют как бы одинаковый ранг: еще не решен вопрос, какие из них будут использованы для синтеза (обучения) системы распознавания, а какие — для контроля. По существу, это файл данных с открытым доступом.

В виртуальной базе ранг данных уже определен и с ВБ непосредственно «работает» система на этапе обучения, контроля и анализа результатов распознавания. Если возникает необходимость пополнения базы данных образами, дополнительно синтезированными из исходных данных или удачно распознанными образами, то и здесь можно не выходить за рамки ВБ. Принципиальное



отличие ВБ от ИБ состоит в том, что данные в ВБ прошли некоторый этап трансформации из данных ИБ (например, сортировку — переупорядочение, переназначение или математическую обработку: каждый образ в ВБ представляет только часть образа исходных данных и т. д.). При этом способ трансформации исходных данных — ключ для структурирования ВБ и ключ/пароль для защиты от несанкционированного доступа к ней.

### Как уменьшить риск ложного распознавания?

При идентификации посетителей или контроля доступа к чему-нибудь (например, к компьютеру, к базе данных или в помещение) с помощью систем распознавания «Visitor Identification» возникает оригинальная проблема неполноты исходной базы данных. Число санкционированных пользователей таких систем обычно невелико (от единиц до 1 – 2 десятков), в то время как число несанкционированных пользователей подобных систем даже трудно представить. Кроме собственно людей, стоящих перед видеокамерой, но не являющихся санкционированными пользователями системы, к их числу, например, можно отнести всевозможные образы, предъявленные для распознавания, но не относящиеся к классу изображений лица живого человека (например, фотографии или рисунки лица человека, изображения ладони, напоминающие контур лица и т. д.), а также другие изображения, издали напоминающие выделенный образ лица. Именно относительно этого числа несанкционированных пользователей любая исходная база данных будет неполной: она всегда будет иметь ограниченное число классов и ограниченное число образов в классе. В принципе, система «Visitor Identification» должна разделять предъявленные ей образы на две группы образов: лица человека (группа 1) и все другие образы (группа 2). При этом в группе 1 система должна уметь отделять (распознавать) лицо живого человека от его фотографии или рисунка.

Один из подходов для решения отмеченной проблемы представлен в [Swe99] и проверен в [Kux99b,d]. Здесь, в частности, предлагается в дополнение к исходной базе данных, позволяющей определять санкционированных пользователей системы распознавания, добавить некоторую базу образов-«ловушек», которая облегчит определение возможных несанкционированных пользователей. Например, к базе данных с лицами были добавлены образы — картины природы, сцены на дорогах, групповые снимки людей, фотографии голов (морд) животных, изображения отдельных предметов, напоминающих овал лица, и т. д.

В рассматриваемых системах «Visitor Identification» [Kux99b,d] эта дополнительная база данных подключалась к ВБ и составляла с ней единую базу. По существу, такие дополнительные образы ВБ были предназначены если не для полного исключения ложно-



го распознавания, то хотя бы для значительного уменьшения его риска.

Обратимся к рисунку 4.1.9а, представляющем ВБ данных системы класса «Visitor Identification» [Кух99b,d]: здесь образы с номерами 1 – 9 представляют пользователей системы распознавания (т. е. группу 1), а образы под номерами 10 – 17 — группу 2. Если в процессе распознавания новый образ был квалифицирован как относящийся к группе 2, то он обозначается (определяется, фиксируется) как образ, не относящийся к санкционированным пользователям.

Итак, для уменьшения риска ложного распознавания посетителей системы, не принадлежащих ИБ и не являющихся санкционированными пользователями, а также образов, не относящихся к лицам (например, рука, рисунок лица и т. д.), в ВБ следует включить образы, принадлежащие группе 2. Естественно, это несколько усложняет процесс распознавания, но главное, увеличивает общий объем данных в ВБ и усложняет ее структуру. Однако ВБ данных может существовать, а может и не существовать в явном виде. Это зависит от класса системы «on-line» (области ее использования или сферы применения), от методов трансформации данных из ИБ в ВБ, от необходимости и частоты извлечения данных из ВБ, от частоты, способа и сложности переналадки системы распознавания на новые данные. В конце концов, это может зависеть и от емкости памяти, предназначенной для реализации системы распознавания.

### Целый образ лица или его часть: что важнее?

Самые простые системы распознавания «Visitor Identification» [Sim00] используют изображения не ИБ данных, а их уменьшенные копии, которые могут не храниться в памяти в явном виде. ВБ в таких системах играет промежуточную роль: яркостные признаки уменьшенных копий исходных изображений можно хранить в матрице признаков, где каждый, например, столбец будет представлять собой признаки одного исходного изображения. Однако ключом доступа к данным в этом случае будет новый формат представления данных: степень и способ их компрессии (уменьшения) и способ конкатенации матрицы исходного образа в вектор.

Анализ самой простой системы распознавания «Visitor Identification», проведенный в разд. 6.1, показал, что на результат распознавания влияют буквально все «внешние» факторы: плохое освещение в поле зрения камеры, несоответствующее удаление объекта распознавания от камеры, «мешающие» элементы на лице, неполное соответствие окружающего лица фона, формы и вариантов предъявленных образов образам в ИБ данных.

Предложенный в [Sim00] способ уменьшения влияния перечисленных нежелательных факторов на результат распознавания



состоит в пополнении ИБ данных различными вариантами образов-эталонов в каждом классе. Именно вариации освещения, ... удаления, ... фона и т. д. в образах-эталонах призваны обеспечить некоторую относительную полноту ИБ данных и благодаря этому улучшить результативность распознавания.

Однако собрать в ИБ все естественные вариации образов-эталонов практически невозможно, да, пожалуй, и не нужно. Это можно сделать искусственно в рамках ВБ данных, используя имеющуюся информацию из различных классов ИБ данных. Покажем это на примере пополнения ВБ новыми образами-эталонами, имеющими различный масштаб. При наличии в ВБ таких образов-эталонов можно практически исключить задачу согласования размеров новых образов в процессе их распознавания, поскольку в ВБ уже находится образ-эталон необходимого размера.

Итак, пусть ИБ данных будет содержать по одному изображению в каждом классе и это изображение включает лицо человека и окружающий его фон. И пусть нам известно, что расстояние между центрами глаз для всей совокупности образов ИБ данных изменяется в пределах от  $S_{\min}$  до  $S_{\max}$ . На основе одного исходного изображения, полученного непосредственно с камеры, сформируем  $Q$  образов-эталонов (образов лиц) с размерами между центрами глаз, определяемыми как:  $S_{\min}, S_{\min} + 1, S_{\min} + 2, \dots, S_{\min} + (Q - 1) = S_{\max}$ .

Описываемый процесс показан на рис. 6.2.1. Здесь из исходного изображения вначале выделяется образ-эталон, для которого затем определяется параметр  $R$  — первоначальное расстояние между центрами глаз.

Используя далее методику, описанную в разд. «Согласование размеров изображений», и значения параметра  $S = S_{\min}, S_{\min} + 1, S_{\min} + 2, \dots, S_{\min} + (Q - 1)$ , получим новую совокупность образов для заданного класса.

Из исходного изображения (см. рис. 6.2.1) получаем, например, 12 новых образов-эталонов (рис. 6.2.2). Над каждым из образов указаны вычисленные значения расстояний между центрами глаз.

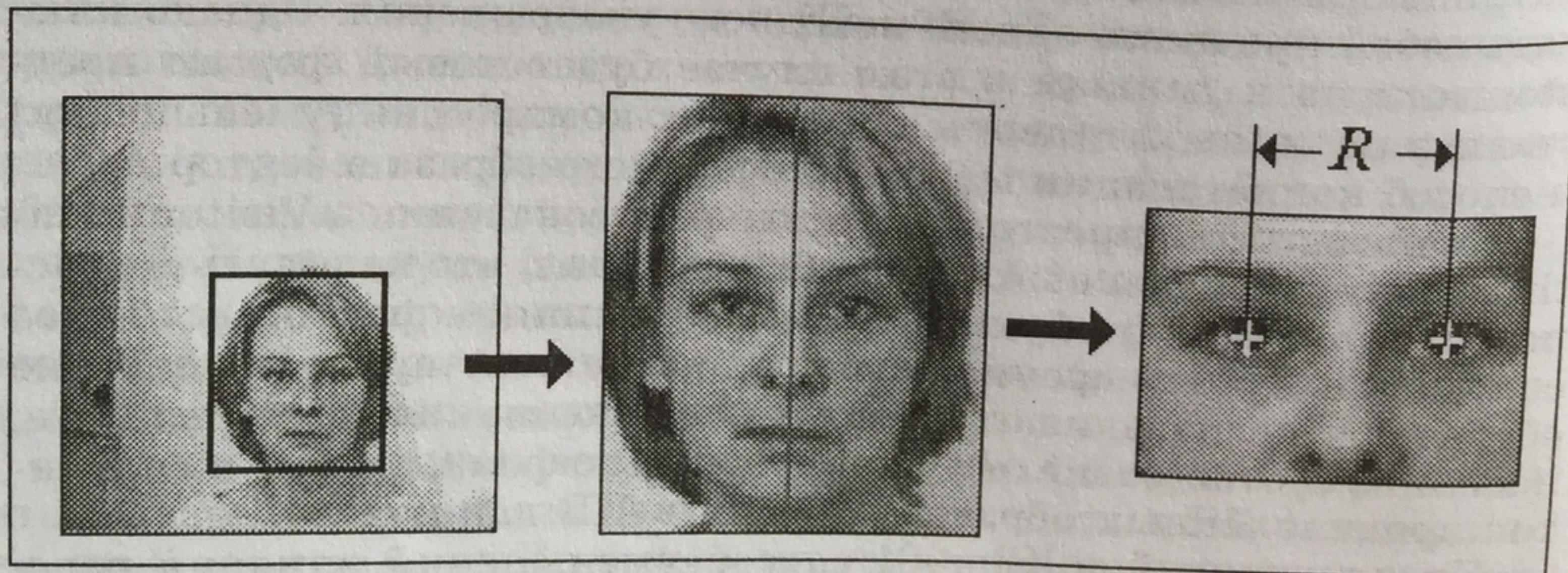


Рис. 6.2.1. Центрирование распознаваемого образа и оценка расстояния  $R$





Рис. 6.2.2. Двенадцать образов-эталонов и вычисленные для них расстояния  $R$

Обратим внимание на один факт — повторение оценок величины  $R$  при некотором заметном даже «на глаз» изменении масштаба образов-эталонов. Иными словами, здесь на лицо результат не совсем точного определения величины  $R$  вследствие трех причин: согласования размеров изображений по критерию  $abs(R - S) \leq \delta$ , выбора величины  $\delta \geq 1$ , имеющей целочисленное значение, и оценки величин  $R$  в числе пикселей, также имеющих целочисленное значение.

К слову сказать, отмеченная ошибка в определении величины  $R$  практически ни на что не влияет в дальнейшем, а каждый новый образ-эталон в ВБ повышает результативность распознавания.

Итак, все полученные образы-эталоны теперь составляют один из классов ВБ. Как теперь, используя только эти до некоторой степени «виртуальные данные», идентифицировать посетителя, если его образ получен в измененных условиях (по отношению к условиям получения изображения в ИБ).

Рассмотрим, например, случай, связанный с изменением фона на изображениях и изменением прически, и оценим принципиальную возможность идентификации такого образа.

Пусть для посетителя рассмотренного выше класса (в данном случае третьего класса в соответствии с ИБ, представленной на рис. 6.1.1) получено непосредственно с камеры новое изображение, а из него выделен образ для распознавания. Этот выделенный образ представлен на рис. 6.2.3а с правой стороны, а с левой стороны на том же рисунке — один из образов, показанных на рис. 6.2.2 и принадлежащих ВБ. Из ВБ выбран именно тот образ, который согласован по размеру с выделенным новым образом.

Из рис. 6.2.3а отчетливо видно, что образы получены в разных условиях (фотографии получены с одной и той же камеры, для одного и того же человека, но в разное время), и вряд ли можно на 100 % утверждать, что на обеих фотографиях одно и то же лицо.



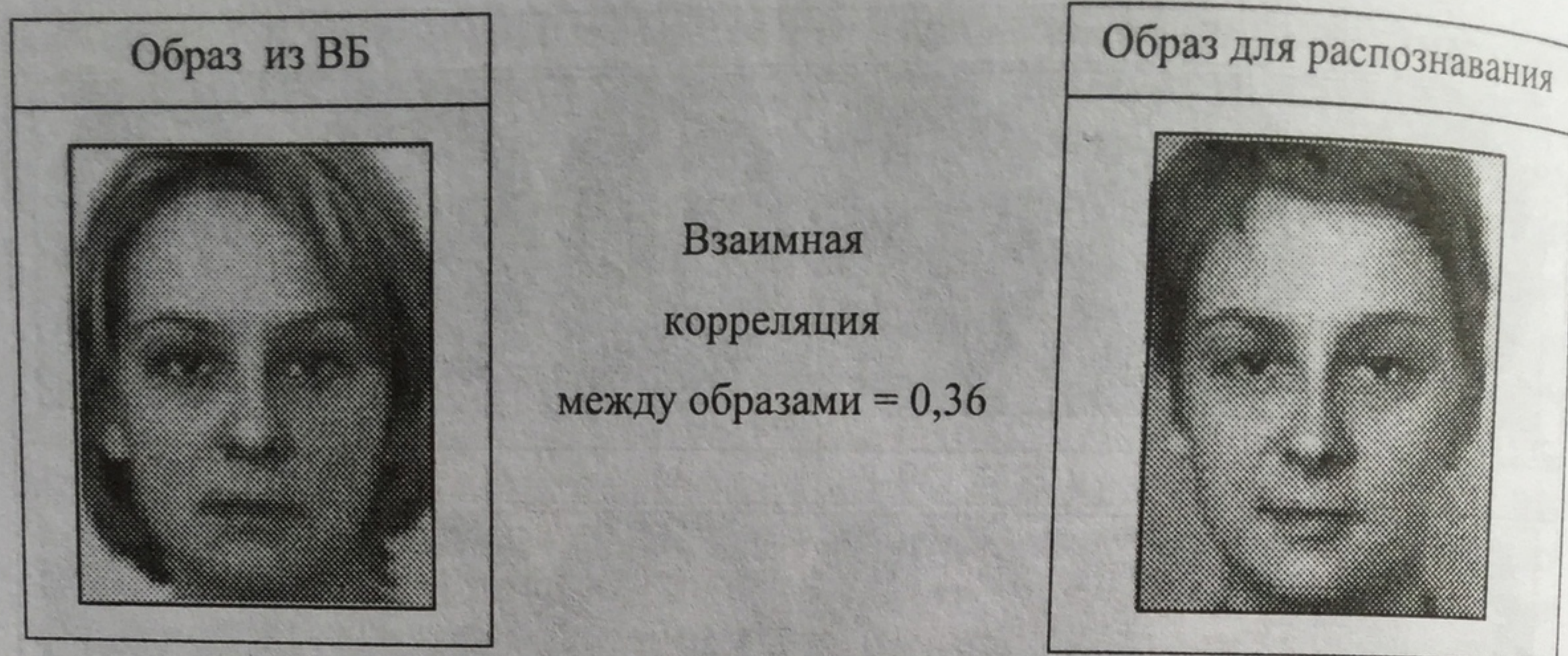


Рис. 6.2.3а. Образ из ВБ и образ, предназначенный для распознавания

Именно поэтому об использовании нормированной функции показателя подобия тут вообще не может идти речи.

Как известно, мерой подобия двух цифровых изображений может служить коэффициент их взаимной корреляции. Однако здесь мы имеем дело со специфическими изображениями лиц человека. Вся известная совокупность таких изображений принадлежит как бы одному гиперклассу, поскольку в центре каждого изображения находится лицо человека, а последнее включает в себя одинаковые для всех изображений части — брови, глаза, нос, рот, овал лица и т. д.

Индивидуальные черты лиц (а именно брови, глаза, нос, рот, овал лица), а также фон/текстура на изображениях в общем случае могут изменяться не столь значительно, чтобы почувствовать их в изменениях амплитудной двухмерной взаимной корреляции: последняя более чувствительна к изменениям в низкочастотной части, которая как раз является более или менее одинаковой для всех изображений лиц. В данном случае лучше использовать не амплитудную корреляцию, а фазовую [Che94], чувствительную ко всем изменениям индивидуальных черт лиц [Kuc99].

Если же оценить меру подобия этих образов с помощью фазовой взаимной корреляции, то в данном случае она составит вели-

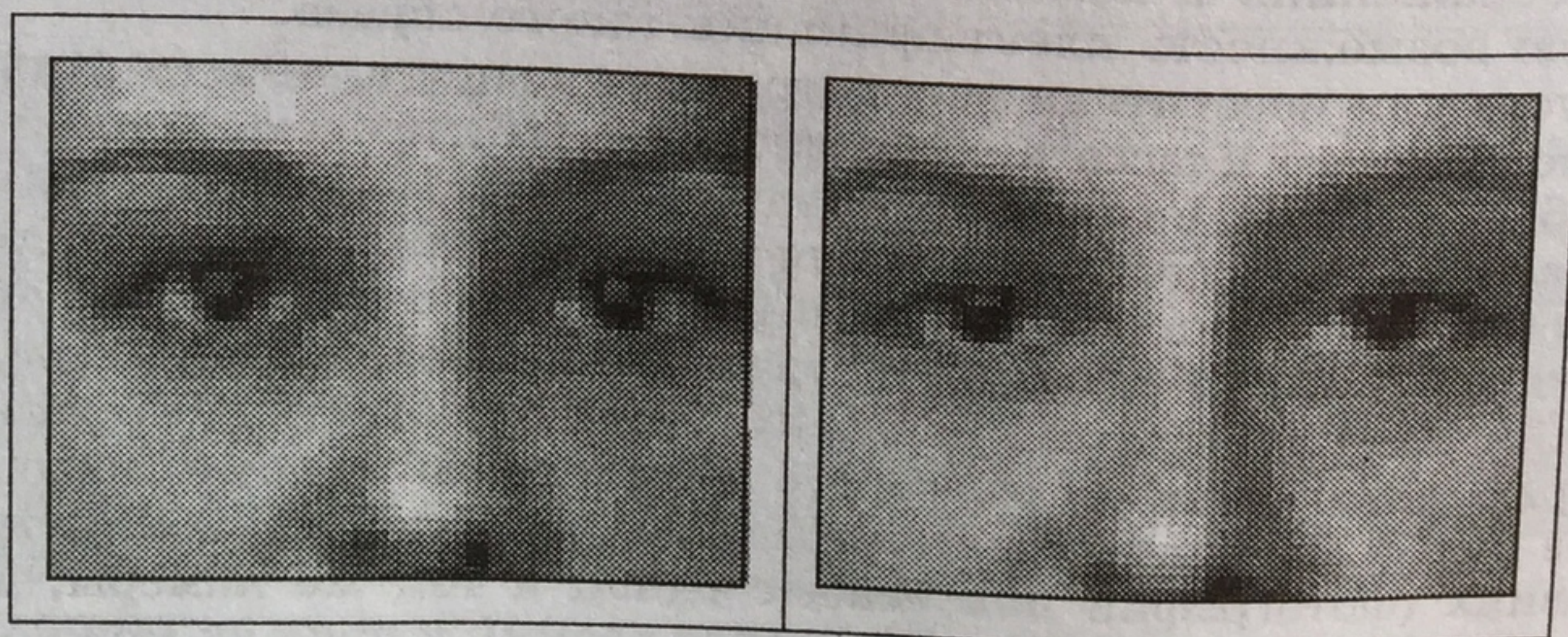


Рис. 6.2.3b. Части образов лиц, представленных на рис. 6.2.3а



чину, примерно равную 0,036, что также не может свидетельствовать о близком подобии этих образов.

В разд. 4.1 мы отмечали, что наиболее информативной и неизменной областью лица является его средняя часть и именно та, которая определяется эталоном № 3 (брови, глаза, нос), представленным на рис. 4.1.2. Для рассматриваемых выше образов эти части лиц показаны на рис. 6.2.3b. Здесь уже, пожалуй, со значительно большей долей уверенности можно утверждать, что на обоих образах представлено лицо одного и того же человека.

Подтверждением тому служит и взаимная (опять же фазовая!) корреляция между выделенными частями образов: здесь она составляет уже величину, равную 0,73, что явно свидетельствует о подобии двух представленных образов.

Заметим, что в обоих случаях мы использовали фазовую корреляцию (см. разд. 3.2) образов, которые перемасштабированы к размеру  $16 \times 16$  пикселей, что соответствует требованиям системы «Visitor Identification» [Sim00].

Анализируя полученные результаты, можно отметить следующее. Низкая корреляция между первыми образами (рис. 6.2.3a) обусловлена в меньшей степени различиями их яркости, фона и некоторыми различиями в масштабе (удалении от камеры в момент съемки), в значительно большей степени — различиями в тонких чертах лица (линия бровей и ресниц, линия переносицы, губ, область рта), в овале лиц, а главное — форме причесок.

Однако и корреляция между вторыми образами (рис. 6.2.3b) не достигла значения, близкого к 1, что также обусловлено различиями в тонких чертах лица (см., например, линию бровей, ресниц и переносицы) и остаточным влиянием части прически (см. правый верхний угол левого образа), а также, возможно, некоторым несоответствием масштабов представления этих образов. Напомним, что в рассматриваемых здесь методах точность согласования размеров двух образов зависит от оценки расстояния между центрами глаз, которая, в свою очередь, определяется с ошибкой  $\delta$ , о чем писалось выше.

И все же во втором случае получено подтверждение того, что образы на рис. 6.2.3 представляют одного и того же человека.

Чтобы убедиться в уникальности этого результата в общем случае, выполним полный эксперимент — вычислим взаимную корреляцию между образами ВБ (см. рис. 6.2.2) и образами этого же класса, предназначенными для идентификации.

Вся совокупность образов, которые можно получить из исходного изображения для всего диапазона изменений параметра  $S$  от 25 до 32 пикселей, показана на рис. 6.2.4. В нашем эксперименте получено 14 новых (можно сказать виртуальных) образов. Все их можно рассматривать как образы, имитирующие приближение посетителя к цифровой фотокамере в системе «Visitor Identification» и, следовательно, как образы, каждый из которых может быть предъявлен для идентификации посетителя.





Рис. 6.2.4. Образы одного и того же класса с разным масштабом их представления, имитирующие приближение посетителя к камере в системе «Visitor Identification»

При этом выделенные части лиц из образов, показанных на рис. 6.2.2 и 6.2.4, показаны на рис. 6.2.5а и 6.2.5b соответственно. Для наилучшего визуального сопоставления образов на этих рисунках они выполнены практически в одном и том же масштабе. Даже при самом беглом просмотре образов на рис. 6.2.5а и 6.2.5b можно отметить их сходство между собой.

А теперь вычислим коэффициенты взаимной фазовой корреляции между соответствующими совокупностями образов. Полученные значения взаимной корреляции ( $12 \cdot 14 = 168$  значений) представлены 3D-графиками (рис. 6.2.6). Здесь левый график — взаимная корреляция между образами, представленными на рис. 6.2.2 и 6.2.4, а правый — взаимная корреляция между образами, представленными на рис. 6.2.5а и 6.2.5b.

Среднее значение корреляции на левом графике — 0,28, а максимальное значение не превосходит 0,5. Среднее значение корреляции

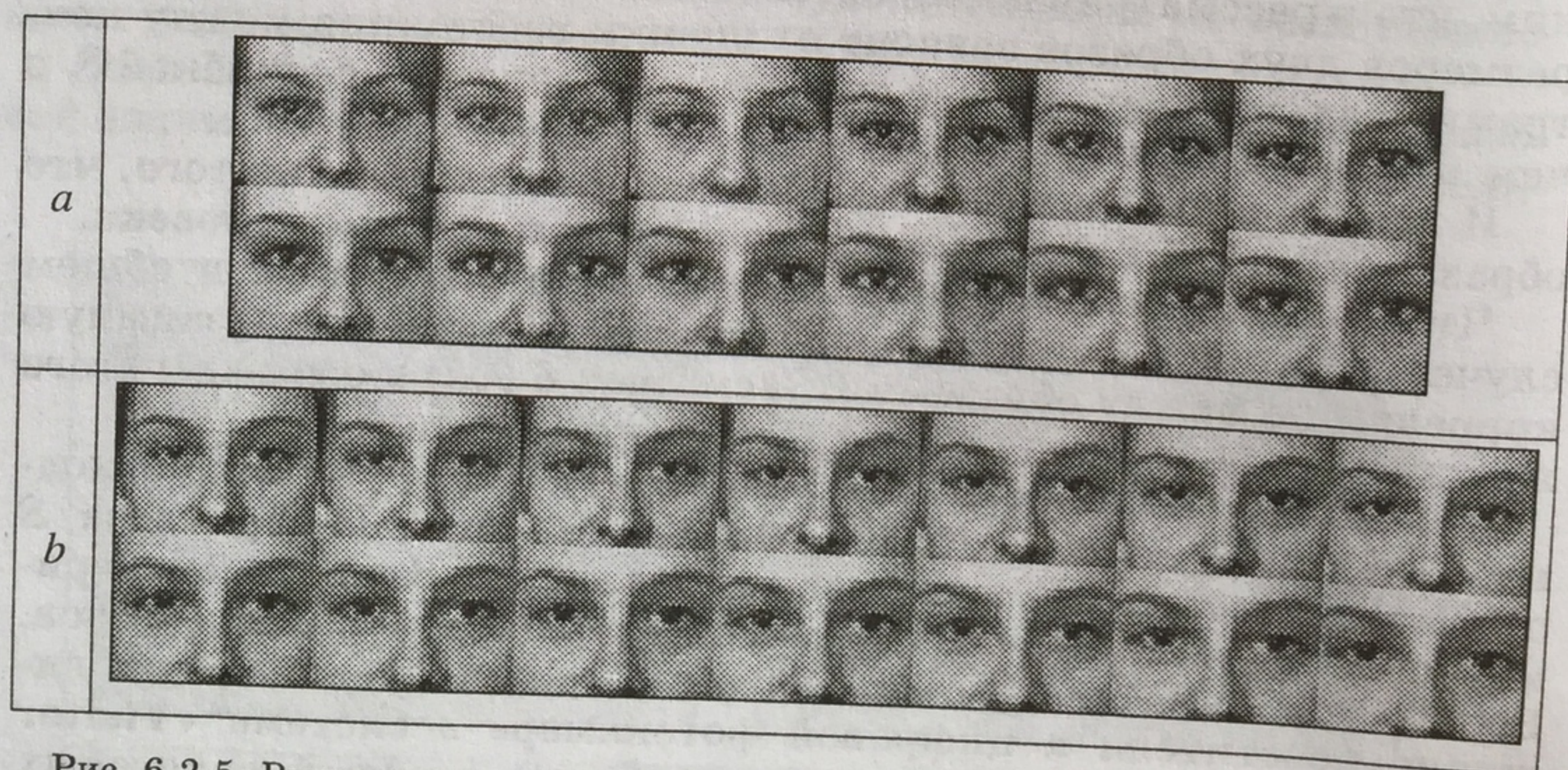


Рис. 6.2.5. Выделенные части лиц из образов, показанных на рисунках 6.2.2 и 6.2.4



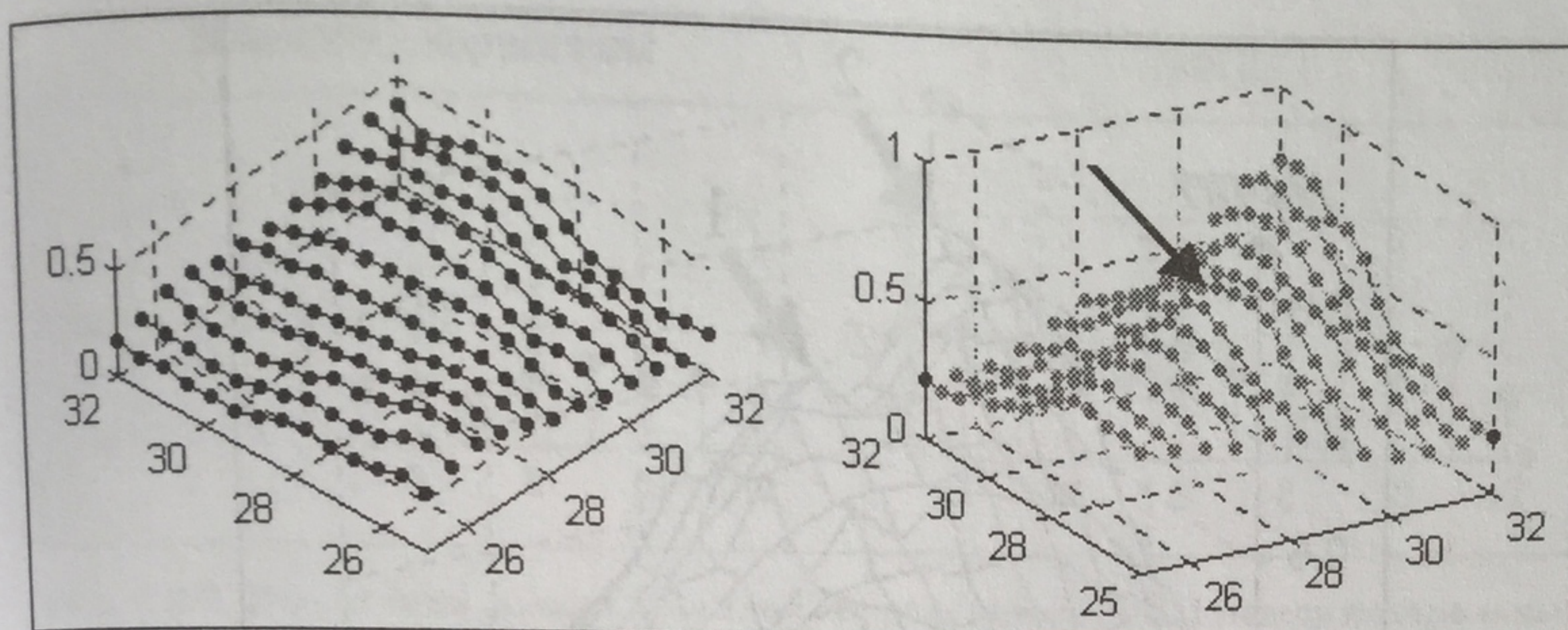


Рис. 6.2.6. Результаты оценки взаимной корреляции между образами ВБ и образами, предъявленными для распознавания, а также между частями этих образов

ляции на правом графике равно 0,46, а максимальное значение значительно превосходит величину 0,5 и составляет 0,75.

При этом обратим внимание на то, что во втором случае корреляция становится больше 0,5 только тогда, когда  $R \geq 28$ , поскольку при значении  $R = 28$  и  $R = 29$  (в данном конкретном случае) происходит более или менее точное согласование размеров обоих образов. Именно здесь отмечается максимум корреляции (его положение отмечено на графике стрелкой).

При значениях  $R > 29$  наиболее информационная часть лица занимает всю область распознаваемого образа, а фон и «мешающие» элементы на образах (например, волосы на верхних образах рис. 6.2.5) постепенно исключаются из рассмотрения.

На основании отмеченного можно сделать вывод, что во втором случае подобие образов можно считать полностью доказанным, и, следовательно, образы на рисунках 6.2.5а и 6.2.5b — это лицо одного человека.

Поскольку полученные результаты обсуждаются в приложении к системам класса «Visitor Identification» (в том числе и описанным в [Sim00]), представим также нормированную функцию показателя подобия (НФПП) для данных рис. 6.2.5а и 6.2.5b. Эта функция, содержащая также все 168 значений, представлена в форме 3D-поверхности на рис. 6.2.7. В данном случае среднее значение НФПП равно 0,59, а максимальное — 0,73, что свидетельствует о принадлежности образов на рисунках 6.2.5а и 6.2.5b одному и тому же лицу.

В заключение следует проверить, каковы показатели подобия распознаваемого образа (из класса 3) аналогично выбранным образам из других классов ВБ?

Пусть, например, на рис. 6.2.8 представлены выбранные образы из 12 классов ВБ, а цифрами 1 – 12 обозначены номера классов, которым принадлежат эти образы.

В качестве контрольного образа выберем 10-й из совокупности образов, представленных на рис. 6.2.5b. Именно для этого образа вычис-



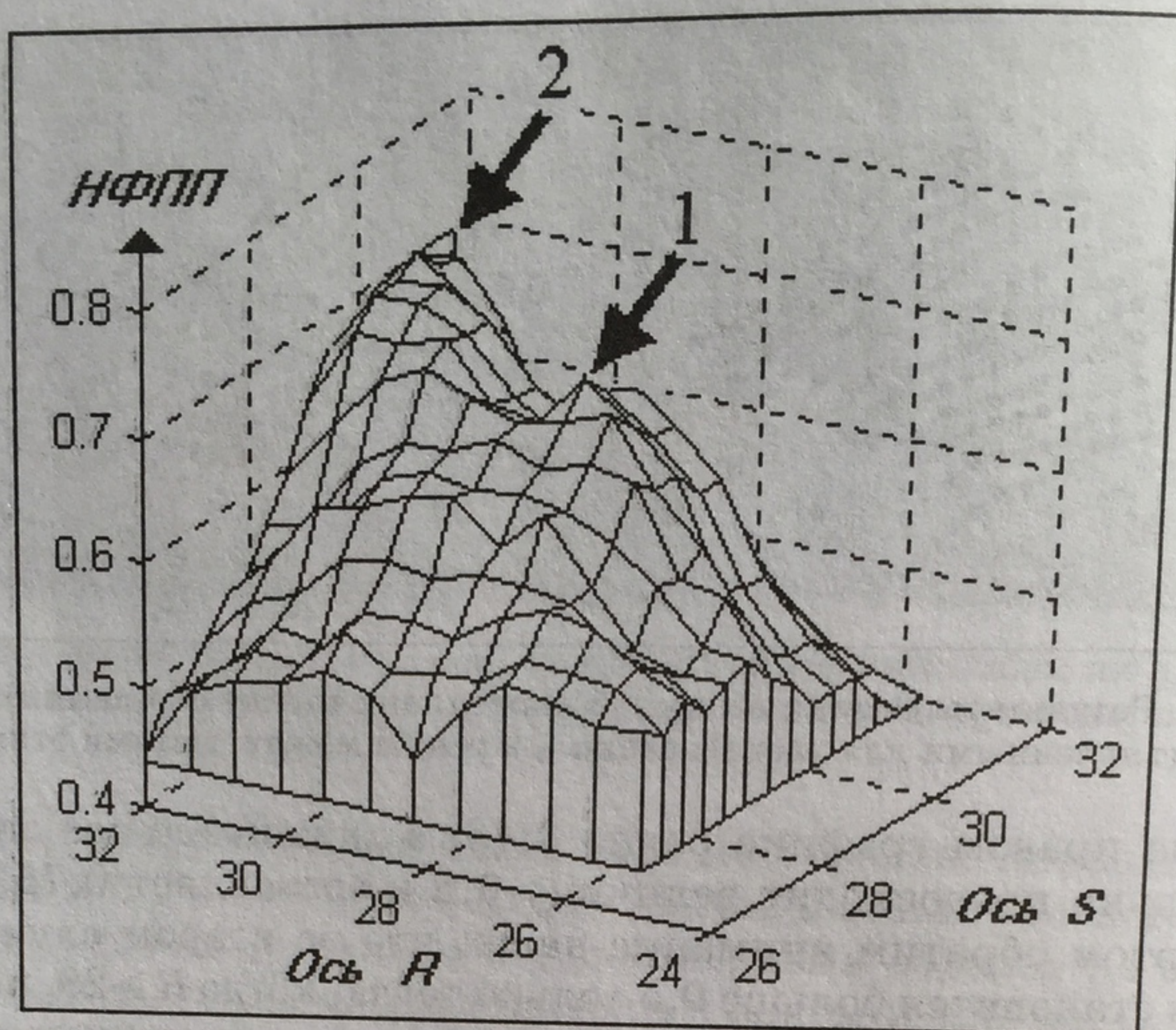


Рис. 6.2.7. Результаты оценки нормированной функции показателя подобия между образами, представленными на рис. 6.2.5а и 6.2.5б

1 — максимум, соответствующий наилучшему согласованию между собой размеров двух образов из совокупностей 6.2.5а и 6.2.5б; 2 — максимум, соответствующий случаю, когда выделенные части лиц занимают всю область в поле рамки образов

ленное расстояние между центрами глаз составило 28 пикселей, что и определило локальный максимум корреляции на рис. 6.2.6.

Вычислим теперь два оценочных показателя — взаимную корреляцию и НФПП между контрольным образом и всеми образами рис. 6.2.8 (включая и 3-й образ). Результаты этих оценок представлены графически на рис. 6.2.9, где на оси Y — соответствующие оценки, а на оси X — номера классов.

Анализируя представленные результаты, отметим, что контрольный образ более всего подобен образу, принадлежащему именно третьему классу. Оба показателя подобия (корреляция и НФПП)

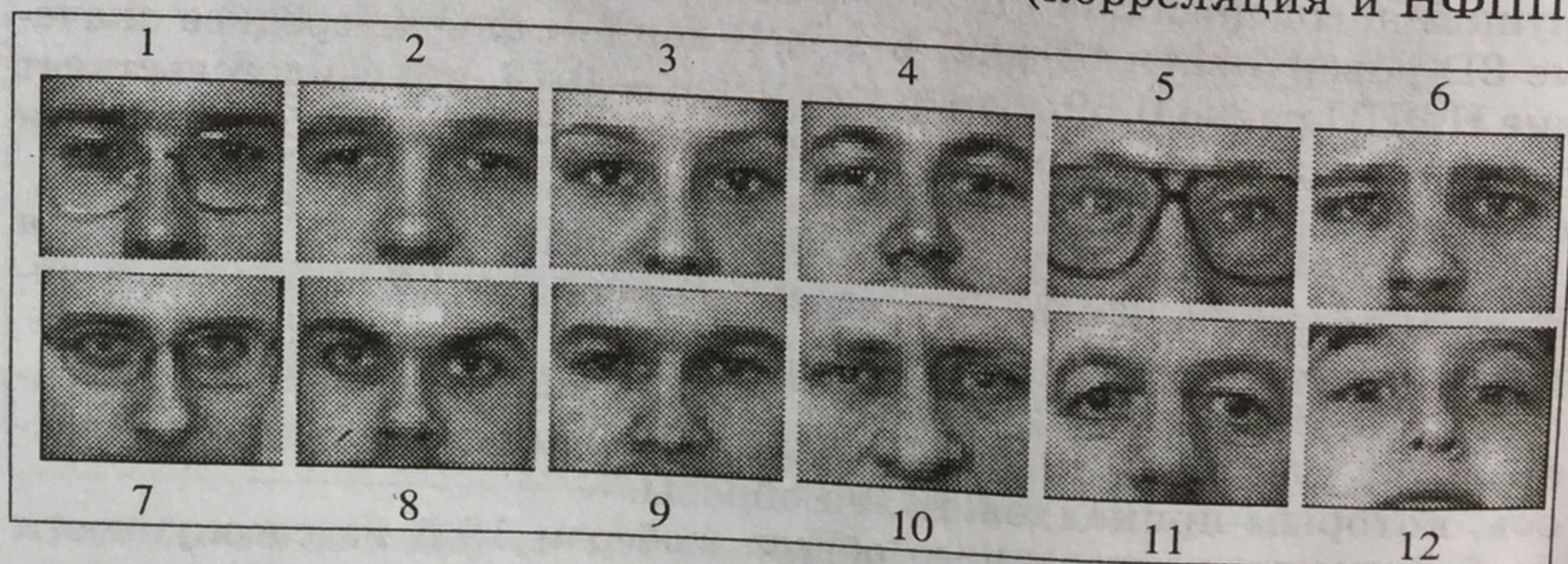


Рис. 6.2.8. Выделенные части лиц из образов 12 классов



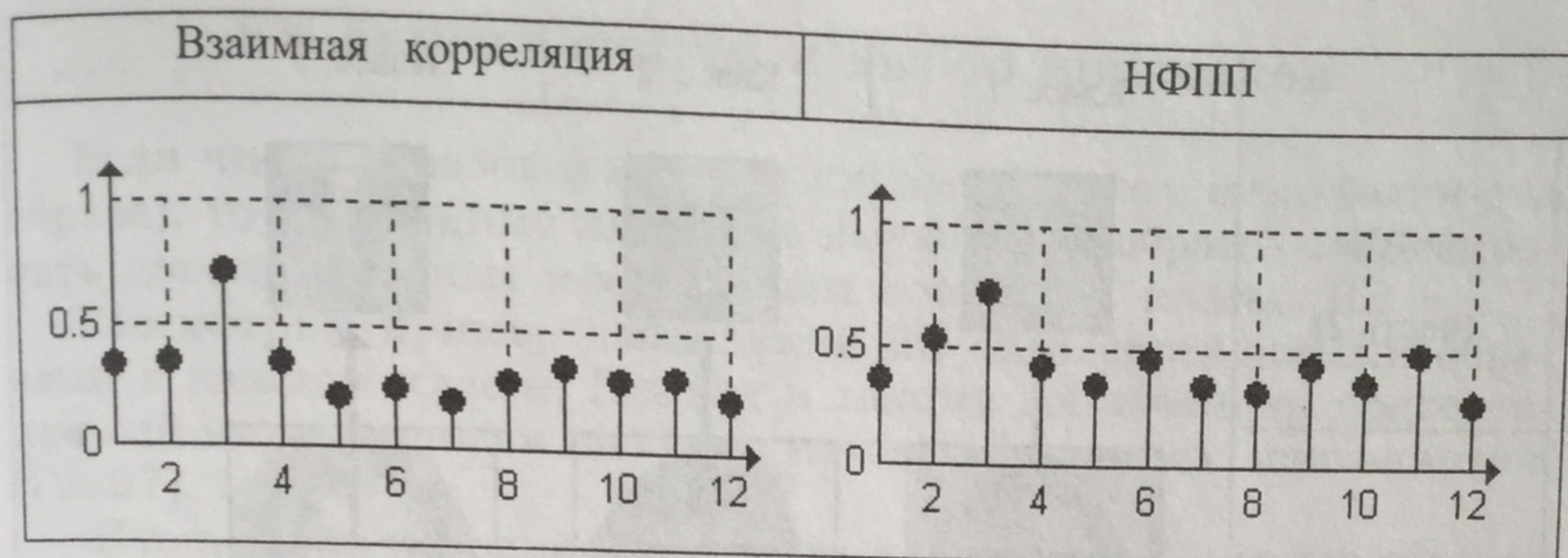


Рис. 6.2.9. Результаты оценки взаимной корреляции и НФПП между контрольным образом и образами, представленными на рис. 6.2.8

здесь превышают 0,7, в то время как показатели подобия с другими образами практически не превосходят величины 0,5.

Итак, выполненные эксперименты показали, что наиболее значимой частью распознаваемого класса изображений является средняя часть лица, определяемая эталоном № 3 (см. рис. 4.2). Образы, полученные на его основе, должны быть включены в состав ВБ и именно по ним может и должна выполняться идентификация личности человека.

Однако, как оказалось, наряду с использованием метрики  $L_0$  и основанной на ней нормированной функции показателя подобия образов в системах «Visitor Identification» для ценки подобия двух выделенных образов можно использовать и взаимную корреляцию. Оценки подобия образов, полученные этими двумя способами, дали практически одинаковые качественные результаты, хотя численные оценки взаимной корреляции были несколько лучше, чем оценки НФПП (оценка НФПП вычислялась для  $\delta = 20$ ).

Таким образом, основной и дополнительный алгоритмы распознавания, входящие в структуру системы распознавания (см. структуру на рис. 6.1.6), могут базироваться на упомянутых выше способах оценки.

В заключение отметим, что дополнение каждого класса ВБ выделенными образами средней части лиц (разумеется, разного масштаба), а также применение основного и дополнительного алгоритмов распознавания образов (фактически двух разных методов оценки подобия) в одной системе «Visitor Identification» создает хорошие предпосылки для улучшения качества распознавания посетителей.

### Построение ВБ системы «Visitor Identification»

С учетом изложенного выше попробуем представить новую структуру ВБ для обсуждаемого класса систем.

Способ построения ВБ для систем класса «Visitor Identification» показан на рис. 6.2.10. Исходная база данных содержит здесь по одному изображению в каждом классе (одно изображение лица



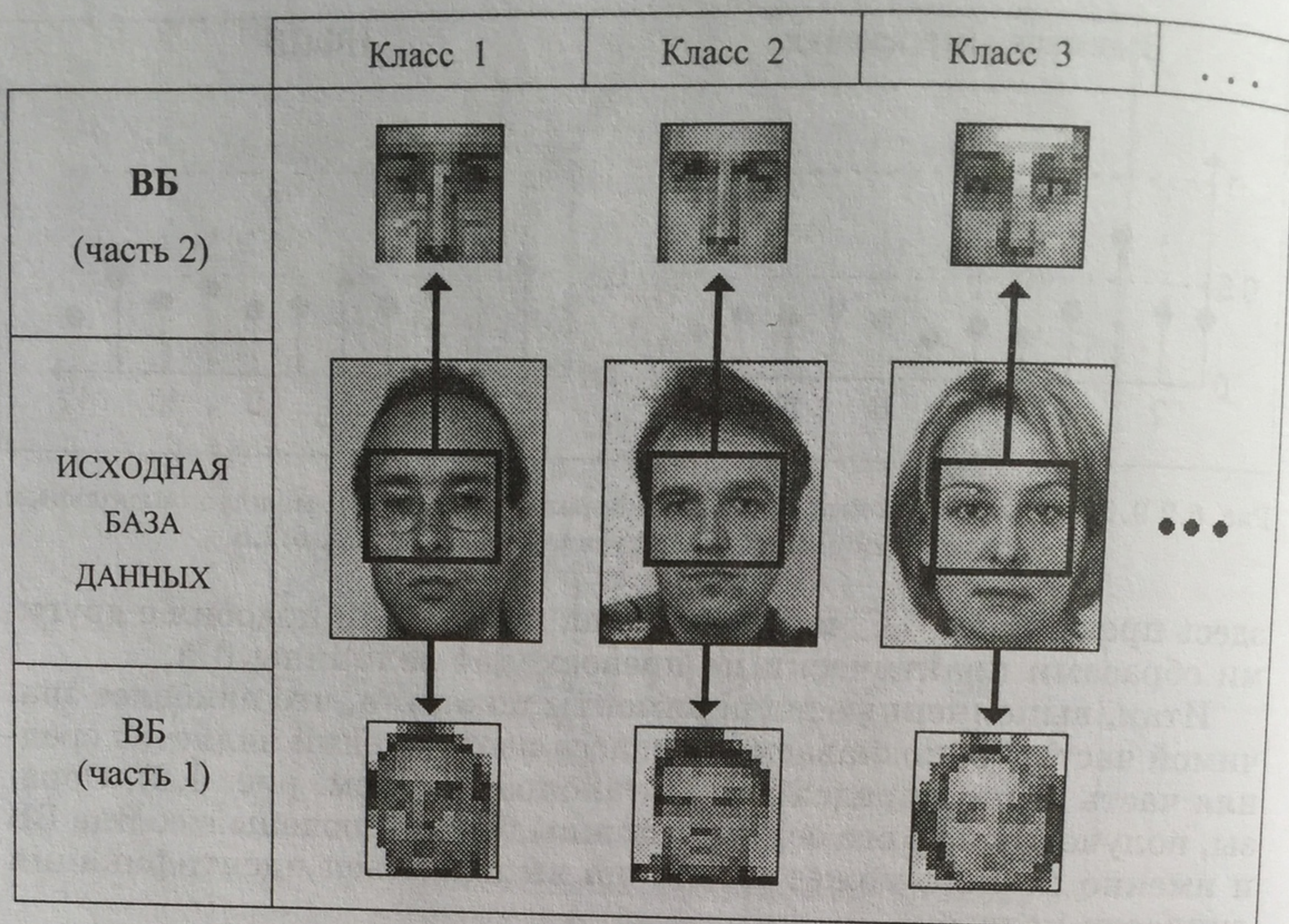


Рис. 6.2.10. Пример построения ВБ из исходных данных

для каждого человека, например, размером  $112 \times 96$  пикселей). Однако каждый класс образов ВБ включает в себя два подобраза: один из них представляет уменьшенную копию целого исходного изображения, а другой — уменьшенную копию изображения средней части лица. Новым элементом по сравнению с системами «Visitor Identification» [Sim00] здесь является второй уровень подобразов в ВБ — уменьшенные копии изображения средней части лица. Как мы помним, выбраны они не случайно: средняя часть лица — не только более информативная, но и менее подверженная изменениям.

Изображение средней части лица имеет размер  $16 \times 16$  пикселей, что позволяет сохранить в нем практически все индивидуальные особенности. При этом ключом/паролем для доступа в ВБ является именно подобраз средней части лица (его размер и способ его использования).

В обсуждаемой системе «Visitor Identification» ВБ присутствует в явном виде. Образы ВБ сохраняются в памяти матрицами соответствующего размера (операция конкатенации в вектор не выполняется), что дополнительно уменьшает проблемы использования метрики  $L_0$ , связанные с циклическими сдвигами собственно лица в поле рамки  $112 \times 96$  пикселей.



## ВБ других систем и выбор признаков

Если число образов в каждом классе слишком мало (один-два образа), то необходимо на основе имеющихся образов синтезировать дополнительные и тем самым пополнить классы ВБ.

Рассмотрим пример дополнения ВБ синтезированными образами в каждом классе. Подход к такому дополнению, представленный ниже, основан на идеях метода «Гусеница», описанного в [Гла97].

Итак, пусть исходная база данных содержит  $K$  классов образов (см. рис. 4.19) и по одному образу в классе. Положим, что в некоторой системе распознавание реализовано на методах PCA + LDA, требующих большого статистического материала. В нашем случае мы имеем дело только с одним образом в классе, а поэтому ВБ необходима для ликвидации информационного дефицита.

Итак, пусть исходные образы имеют размер  $M \times N$  пикселей.

ВБ строится следующим способом:

- в каждом классе образов ИБ с использованием эталона № 3 (см. рис. 4.2) и процедуры выделения средней части лиц определяются центры глаз, вычисляется расстояние между ними, определяются ось симметрии лица и положение линии глаз на исходном изображении;

- далее из каждого образа ИБ выбирается  $Q$  подобразов, представляющих среднюю часть лица, путем сканирования исходного изображения по изменяемым относительно оси симметрии и линии глаз значениям координат.

На рис. 6.2.11 обозначены два варианта (начальный и конечный) выбора подобразов на изображении лица. В данном случае, например, рамка имеет размер  $M_1 \times N_1 = 70 \times 100$  пикселей для исходного изображения размером  $M \times N = 224 \times 184$  пикселя.

Один из вариантов всех таких подобразов представлен на рис. 6.2.12. В данном примере в результате сканирования

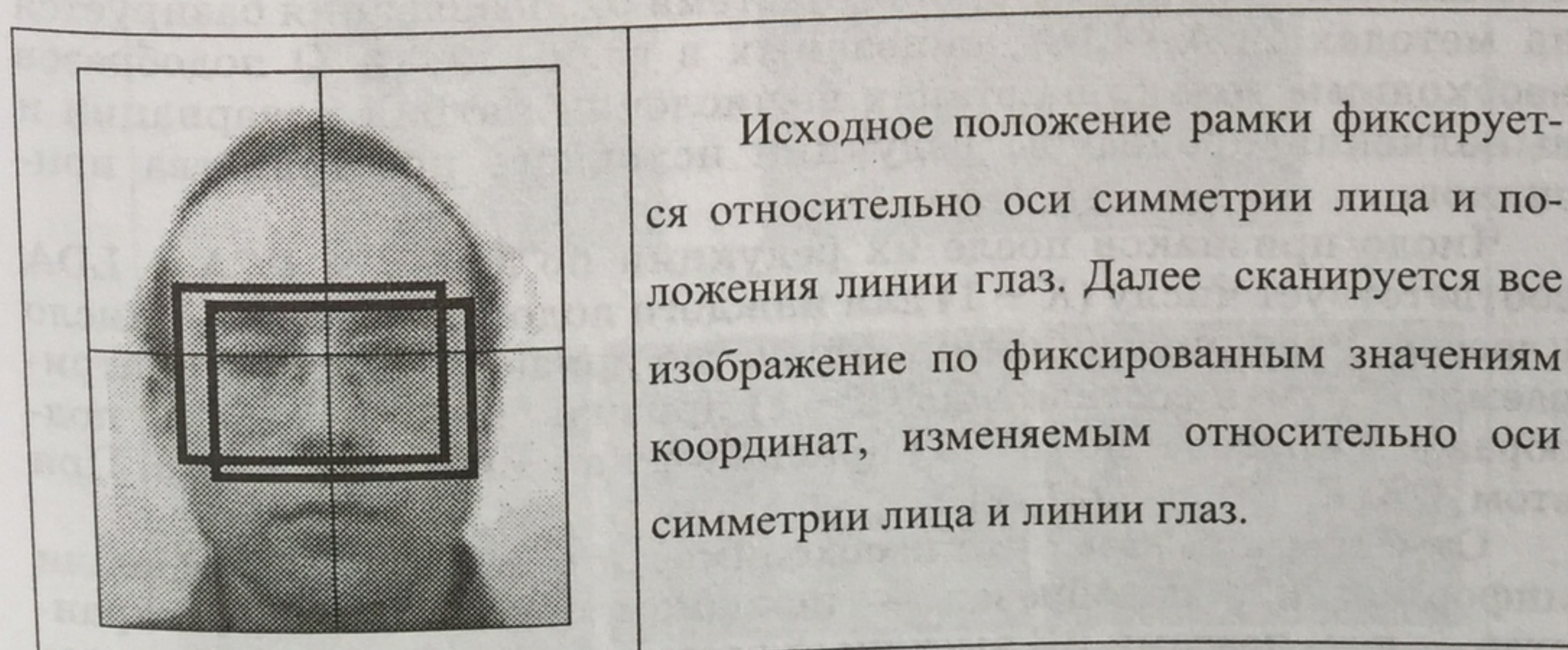


Рис. 6.2.11. Варианты выбора подобразов из исходного изображения лица



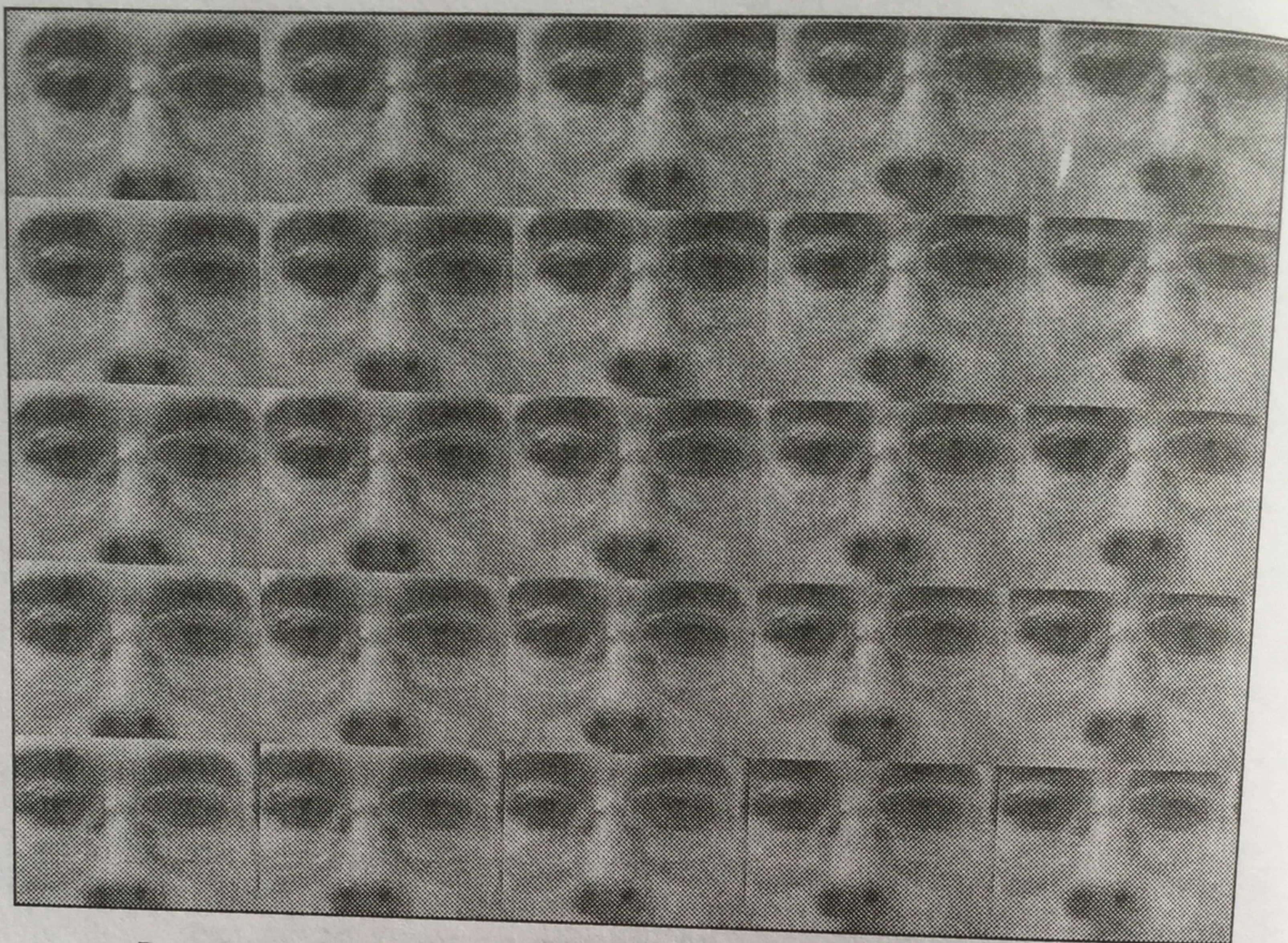


Рис. 6.2.12. Образы ВБ, полученные из образа первого класса ИБ

исходного изображения получается 25 подобразов, каждый размером  $M_1 \times N_1$  пикселей ( $70 \times 100$  пикселей).

Исходные признаки определяются при этом значениями пикселей подобразов, что составляет  $M_1 \times N_1$ , т. е. ( $70 \times 100 = 7000$  признаков) для каждого подобраза. А все признаковое пространство для одного класса ВБ определяется  $QM_1N_1$  признаками, в данном случае  $25 \times 70 \times 100 = 175000$  признаками.

Если учесть, что в ИБ имеется больше одного класса данных, то сохранение этой информации потребовало бы значительных объемов памяти. С другой стороны, система распознавания базируется на методах PCA + LDA, описанных в гл. 5, и эти  $Q$  подобразов необходимы только на этапах вычисления матриц ковариации и выполнения процедуры редукции исходного пространства признаков.

Число признаков после их редукции по методам PCA + LDA соответствует числу  $(K - 1)$  для каждого подобраза, где  $K$  — число классов. Редуцированное пространство признаков для рассматриваемого случая составит  $QK(K - 1)$  признаков:  $K$  классов,  $Q$  подобразов в классе и по  $(K - 1)$  признаков на каждый подобраз. При этом  $QK(K - 1) \ll QM_1N_1$ .

Отметим при этом, что необходимости в длительном хранении информации о подобразах — исходном признаковом пространстве — нет, поэтому ВБ выступает здесь в неявном виде. Для дальнейшего использования при распознавании в памяти сохраняются



$QK(K - 1)$  признаков редуцированного пространства, матрицы редукции пространства признаков, параметры системы и информация о расстоянии между центрами глаз в каждом классе образов.

Итак, главный эффект от дополнения ВБ выбранными образами в каждом классе состоит в возможности применения в системе распознавания мощного математического аппарата методов PCA и LDA при «статистической неполноте» исходных данных.

Теперь остановимся на новой проблеме, также присущей системам распознавания, которые работают в режиме «on-line» с камерой (вариации освещения).

Как бы мы ни старались пополнить ИБ изображениями с различной степенью яркости или как бы ни старались нормализовать яркость исходных изображений, нам не удастся полностью элиминировать яркостную компоненту в пространстве редуцированных признаков. В процессе распознавания нового образа эта компонента будет играть определенную роль, а значит, результат распознавания будет зависеть от яркости распознаваемого образа. В случаях, когда яркостная признаковая компонента приобретает главную роль, а форма и детали лица — второстепенную, необходимо перейти к категории других признаков, например, к нормированным градиентам яркости.

На рис. 6.2.13 представлено исходное изображение лица 1, средняя часть лица, соответствующая фрагменту, выделенному рамкой 2, а также варианты градиентов яркости, отмеченные цифрами 3 — 5.

Образ 3 соответствует суммарному градиенту яркости по оси X и Y, образ 4 — градиенту яркости по оси X, а образ 5 является гра-

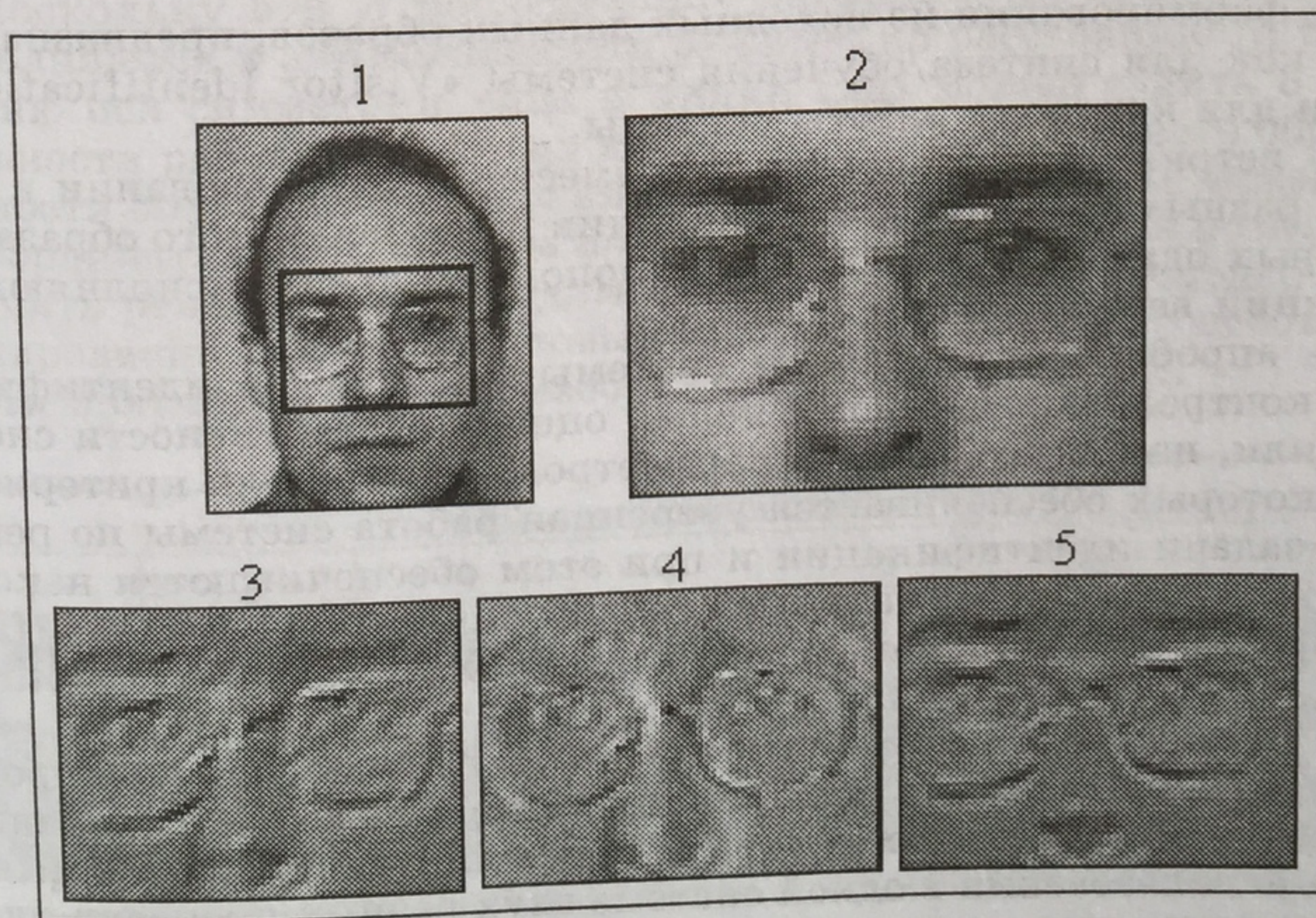


Рис. 6.2.13. Исходный образ, выделенный фрагмент и его градиенты



диентом яркости по оси  $Y$ . Здесь представлены градиенты, нормированные по (4.2). Перед тем как из них выбирать признаки — значения пикселей, нормированное изображение градиентов центрируется на среднее значение изображения.

Главный эффект такого выбора признаков — практическое элиминирование яркостной компоненты в исходном признаковом пространстве и окончательное ее исключение в редуцированном пространстве признаков. Следствием этого является независимость качества распознавания от условий освещенности лица перед камерой.

### 6.3. СИНТЕЗ СИСТЕМЫ «VISITOR IDENTIFICATION»

Представим несколько новых вариантов построения систем класса «Visitor Identification», в которых предусмотрены различные способы повышения точности идентификации посетителя, а также предусмотрена верификация результата его идентификации/распознавания. В этих вариантах будут реализованы новые подходы как в процессе синтеза собственно системы «Visitor Identification», так и в процессе решения системой основной задачи — идентификации/распознавания посетителя.

1. Варианты построения системы при ее синтезе основаны на следующих подходах [Kuc99b, Kuc00a, Kuc00d, Kux99b, Kux99d]:

- дополнении ВБ совокупностью данных, представляющих образы соответствующего класса ИБ, но отличающихся друг от друга размерами (что имитирует приближение/удаление посетителя к камере в момент сеанса с ним).

- формировании из исходных данных образов, предназначенных как для синтеза/обучения системы «Visitor Identification», так и для контроля работы системы.

- встраивании в систему и совместном использовании в ней двух разных процедур классификации распознаваемого образа, из которых одна основная, а другая дополнительная, исполняющая функции верификатора.

- апробировании работы системы на примере идентификации контрольных образов с целью оценки эффективности системы (или, например, оценки параметров, порогов или критериев), при которых обеспечивается уверенная работа системы по решению задачи идентификации и при этом обеспечиваются некоторые требуемые (минимальные) уровни ошибочных отказов (false reject rate — FRR) и ошибочных подтверждений (false acceptance rate — FAR [Hon98, Фил00]).

2. Варианты построения системы в части, реализующей процедуры распознавания и верификации результата распознавания, основаны на следующих подходах [Kuc99b, Kuc00a, Kuc00d]:

- использовании в одной системе двух разных процедур классификации распознаваемого образа, из которых одна основная



Рис. 6.3.



(имеющая высший приоритет), а другая дополнительная, исполняющая функции верификатора;

— выполнении распознавания не только по исходному (всему) образу лица, но и по его наиболее информативной части (аналогичной, например, эталонам № 2 и 3 из рис. 6.2.2), что призвано обеспечить более высокую точность идентификации;

— верификации первоначального результата распознавания за счет контрольной проверки — распознавания дополнительных образов (двух или четырех), полученных непосредственно из исходного изображения путем его перемасштабирования, и использования мажоритарного принципа оценки результата распознавания (по двум из трех или по трем из пяти);

— использовании PIN, вводимых пользователем (посетителем) в систему по ее запросу.

**Рассмотрим конкретный пример.** Пусть проектируемая система «Visitor Identification» предназначена для идентификации семи пользователей (семи санкционированных посетителей), фотографии которых представлены на рис. 6.3.1. Для каждого из них в ИБ данных имеется только по одному изображению, полученному в условиях, в которых будет работать система.

Процесс синтеза данных ВБ на примере образов первого класса представлен на рис. 6.3.2.

Здесь из единственного изображения (отмечено цифрой 1), полученного непосредственно с видеокамеры, выделен первоначальный образ лица (отмечен цифрой 2). На этом образе определены координаты центров глаз и расстояние  $R$  между ними. Далее на основе этих оценок определяются ось симметрии лица и положение линии глаз (см. белые линии на выделенном образе 3).

Поскольку эти и последующие операции на изображениях выполняются в автоматическом режиме, по расстоянию  $R$ , положению оси симметрии лица и линии глаз можно судить о правильности работы алгоритма выделения первоначального образа, точности определения координат центров глаз, а также о разнице в координатах центров глаз по оси  $Y$ . Эта информация позволяет принять решение о необходимости использования процедур элиминирования наклонов головы (в плоскости  $XY$ ) или о необходимости получения другого изображения с видеокамеры.

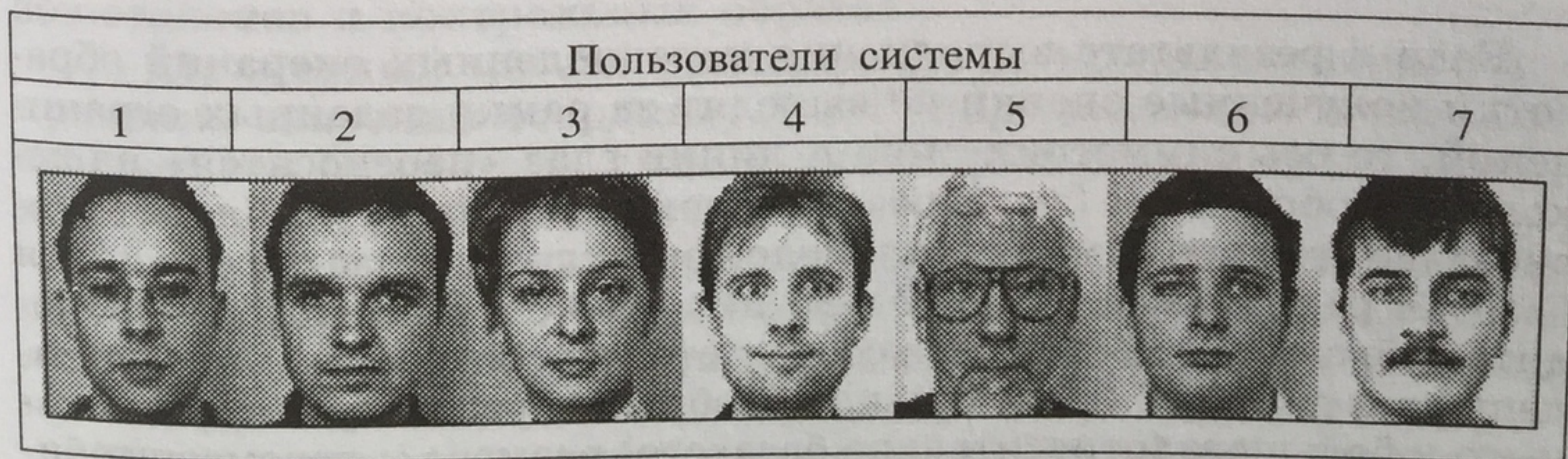


Рис. 6.3.1. Семь санкционированных пользователей системы «Visitor Identification»



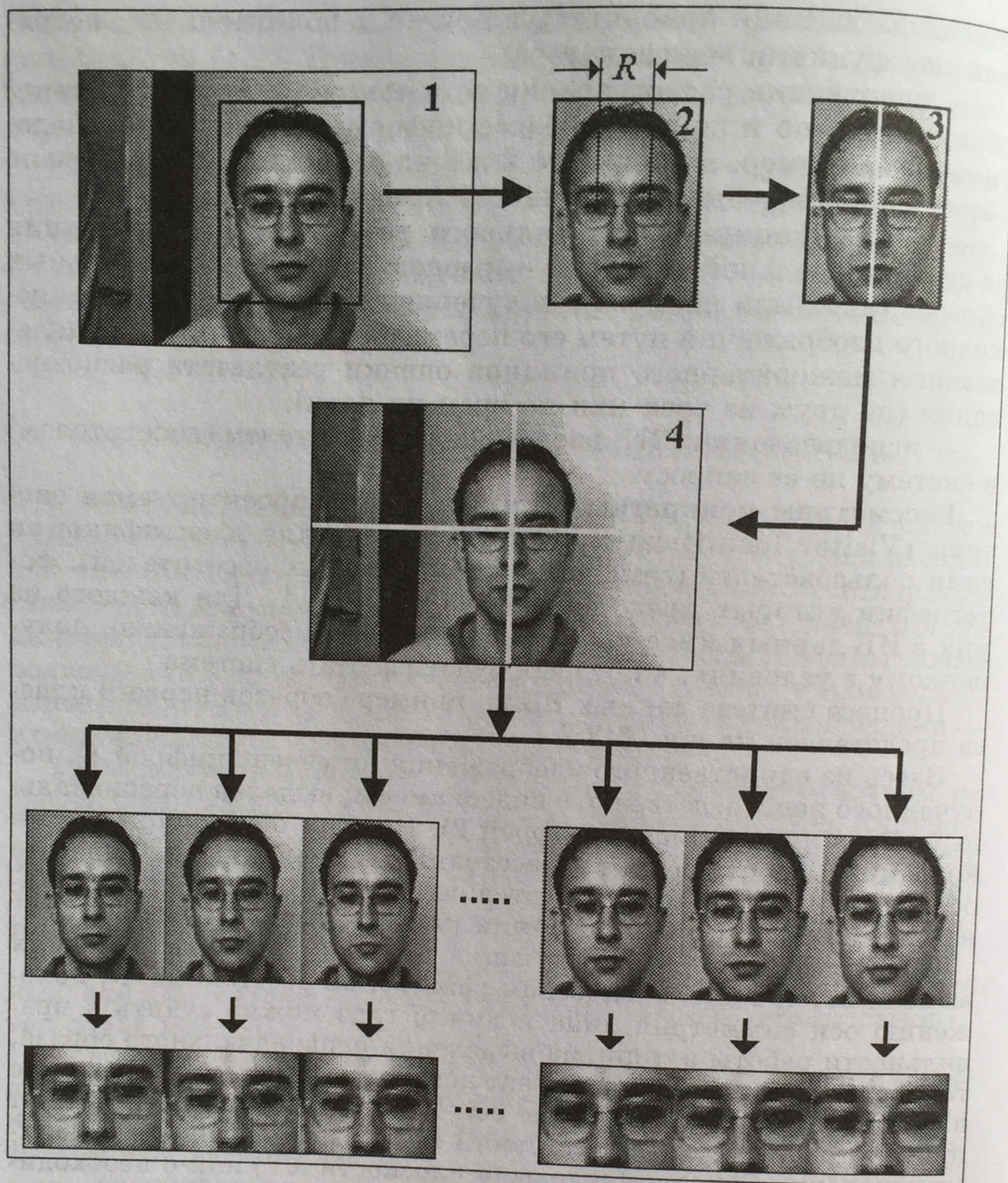


Рис. 6.3.2. Процесс синтеза данных ВБ на примере исходного образа первого класса

Если в результате выполнения перечисленных операций обработки полученные оценки не выходят за рамки заданных ограничений, то ось симметрии лица и линия глаз «переносятся» на исходное изображение (см. образ, отмеченный цифрой 4). Далее относительно этой оси симметрии и положения линии глаз определяется базовая рамка размером  $112 \times 96$  пикселей, в поле которой находится образ лица. Выбирая в соответствии с алгоритмом, представленным в разд. 6.2.2, несколько подобных образов с рамками меньшего и большего (относительно базового) размера и перемасштабируя их к одному и тому же размеру ( $112 \times 96$  пикселей), можно



получить  $L$  образов лица с разными, но монотонно изменяющимися значениями величины  $R$ . В полученном ряду образов величина  $R$  меняется в пределах заданных значений от  $R_{\min}$  до  $R_{\max}$ . Этот ряд имитирует приближение/удаление посетителя к камере в момент сеанса с ним. Полученные при этом образы-эталоны используются далее для получения из них соответствующих подобразов. Описанный процесс представлен схематически в нижней части рис. 6.3.2.

Контрольные образы получают аналогичным способом. Один из образов определяется базовой рамкой на исходном изображении, и этот образ точно соответствует одному из  $L$  упомянутых выше. Два других образа определяются меньшим и большим (относительно базового) размером рамки. При этом изменения размеров рамки фиксированы некоторой величиной  $\pm\Delta$  и в общем случае не связаны с согласованием размеров по параметру  $R$ .

Естественно, что описанный выше процесс синтеза данных ВБ повторяется для всех  $K$  классов образов. В рассматриваемом нами случае он повторился бы 7 раз.

Далее в соответствии с методологией построения систем «Visitor Identification» [Кух99, Кус00], образы-эталоны и контрольные образы, а также соответствующие им подобразы перемасштабируются к размеру  $16 \times 16$  пикселей. Каждый такой образ преобразуется затем (с помощью операции конкатенации) в вектор, содержащий 256 элементов, и записывается в память под номером своего класса.

По окончании процедуры синтеза новых образов в памяти системы «Visitor Identification» содержатся следующие данные:

- 1) параметры системы —  $\{K, L_k, Q\}$ ,  
где  $K$  — число классов;  $L_k$  — число образов-эталонов в  $k$ -м классе,  $k = 1, 2, \dots, K$ ;  $Q = \max\{L_1, L_2, L_3, \dots, L_K\}$ ;
- 2) значения размеров всех используемых масок;
- 3) значения порогов  $\delta$ , используемых в метрике  $L_0$ ;
- 4)  $K$  исходных изображений, полученных непосредственно с видеокамер;
- 5)  $K$  изображений лиц пользователей (образов-эталонов размером  $112 \times 96$  пикселей);
- 6) двумерный массив размером  $(Q + 3) \times K$ , предназначенный для хранения информации о значениях величин  $R$  для всех образов-эталонов и контрольных образов;
- 7) двухмерный массив размером  $2 \times K$ , предназначенный для хранения значений координат  $X$  и  $Y$ , определяющих точку пересечения оси симметрии лица и линии глаз на исходных изображениях;
- 8) два трехмерных массива размером  $256 \times Q \times K$ , определяющих образы-эталоны и соответствующие им подобразы (в случаях, когда  $L_k < Q$ , неиспользуемые столбцы массивов заполняются нулями);
- 9) два трехмерных массива размером  $256 \times 3 \times K$ , определяющих контрольные образы и соответствующие им подобразы;



Прежде чем приступить к апробированию работы системы на примере идентификации контрольных образов, представим ее структуру (рис. 6.3.3). При этом за основу примем структуру, представленную на рис. 6.1.6 и описанную в разд. 6.1. Мету подобия между образами-эталоны и контрольным образом (или образами-эталоны и каждым новым образом) будем определять двумя методами. Первый из них — основной — базируется на взаимной корреляции, а второй — дополнительный — на основе метрики  $L_0$ .

На структурной схеме системы цифрами 1 и 2 обозначены посетитель системы «Visitor Identification» и исходное изображение, полученное с камеры. Цифрами 3a, 3b и 3c отмечены три выделенных из исходного изображения образа лица, предназначенных для идентификации (в режиме тестирования системы это будут контрольные образы). Причем образ 3b получен непосредственно из исходного изображения (без изменений его масштаба), а образы 3a и 3c соответствуют удалению лица от видеокамеры и приближению лица к видеокамере. В связи с этим на образе 3a изображение лица будет уменьшенным, на образе 3c — увеличенным по отношению к образу 3b.

Цифрами 3' отмечены перемасштабированные к размеру  $16 \times 16$  пикселей образы 3a, 3b и 3c и соответствующие им подобразы — части лиц, представленные в форме векторов. Именно они сравниваются с образами и подобразами 4' всех классов данных из ВБ, также представленными в форме векторов.

Цифрами 5 и 6 отмечены две группы результатов вычисления нормированного показателя подобия, основанного на метрике  $L_0$  — для целых образов лиц и их подобразов соответственно. Для  $K$  классов и  $L_k$  образов в каждом классе, в каждой группе получаем по  $KL_k$  оценок взаимной корреляции на каждый из образов 3'.

Цифрами 7 и 8 обозначены две группы результатов вычисления взаимной корреляции — для целых образов лиц и их подобразов соответственно. Здесь также получаем по  $KL_k$  оценок показателя подобия в каждой группе и для каждого из образов 3'.

Цифра 9 определяет окончание анализа результата классификации нового образа и принятие решения. Далее идет отображение результата в случае успешной идентификации посетителя (линия с цифрой 10) и/или запись результата в протокол сеанса с посетителем (линия с цифрой 12). Цифра 11 — путь повторного запуска видеокамеры в случае отрицательного результата распознавания.

Положим, что основной метод классификации в данной системе основан на взаимной корреляции. Дополнительным методом, выполняющим функции верификации, будет в таком случае метод классификации на основе метрики  $L_0$ .

В случаях, когда результаты классификации по основному алгоритму и верификации по дополнительному, являются одинаковыми, распознаваемый образ считается идентифицированным.

В случаях, когда результаты классификации по основному алгоритму и верификации по дополнительному, разные, распознаваемый образ считается неидентифицированным.







Рассмотрим теперь подробнее работу блока анализа результатов классификации и принятия решений.

Будем считать, что блоком «Определение меры подобия образов на основе корреляции» вычислены коэффициенты взаимной корреляции одного образа группы 3' (конкретно образа 3а, 3б или 3с) со всеми образами группы 4' (с каждым образом и в каждом классе). Из всей совокупности полученных оценок корреляции выбираем только максимальные значения. При этом для каждого из образов группы 3' (т. е. отдельно для образа 3а, 3б или 3с) и соответствующих им подобразов получаем следующие оценки:

- максимум корреляции для образа целого лица  $Ko(i)$ ;
- максимум корреляции для подобраза лица —  $Kp(j)$ ;
- максимум средней корреляции в классе для образа целого лица  $So(l)$ ;

- максимум средней корреляции в классе для подобраза лица  $Sp(m)$ , где  $i, j, l$  и  $m$  — номера классов (в перечислении их по ВБ), которым принадлежат максимумы.

Будем также считать, что блоком «Определение меры подобия образов на основе метрики  $L_0$ » получены оценки нормированной функции показателя подобия одного образа группы 3' (конкретно образа 3а, 3б или 3с) со всеми образами группы 4' (с каждым образом и в каждом классе). Из всей совокупности полученных оценок показателя подобия выбираем только максимальные значения. При этом для каждого из образов группы 3' (то есть отдельно для образа 3а, 3б или 3с) и соответствующих им подобразов получаем следующие оценки:

- максимум показателя подобия для образа целого лица  $Fo(h)$ ;
- максимум показателя подобия для подобраза лица  $Fp(g)$ , где  $h$  и  $g$  — номера классов (в перечислении их по ВБ), которым принадлежат максимумы функции показателя подобия.

На основе практических исследований, выполненных на большом объеме экспериментальных данных в соответствии с методикой, рассмотренной в раз. 6.3, можно утверждать, что наибольший интерес для окончательного принятия решений об идентификации представляют следующие оценки:  $Kp + Sp$ ,  $So + Sp$ ,  $Fo$  и  $Fp$ .

Далее будем исходить из того, что вся совокупность оценок  $\{Kp + Sp, So + Sp, Fo \text{ и } Fp\}$  будет определять принадлежность каждого распознаваемого образа (3а, 3б или 3с) к одному и тому же классу только если выполняется одно из следующих условий:

$$(Kp + Sp) > P_1 \text{ при } j = g; \quad (6.3.1)$$

$$(So + Sp) > P_2 \text{ при } i = j; \quad (6.3.2)$$

$$(So + Sp) > P_3 \text{ при } j = (h + g)/2; \quad (6.3.3)$$

$$(So + Sp) > P_4 \text{ при } (i + j + l + m + h + g)/6 = i, \quad (6.3.4)$$

где  $P_1 — P_4$  — некоторые установленные пороги, причем  $P_1 > P_2 > P_3 > P_4$ .



Из приведенных условий видно, что первое из них и последние два используют все вычисленные оценки, и здесь, таким образом, для подтверждения распознанного класса используется процедура верификации.

А теперь при окончательном принятии решения используем мажоритарный принцип. Напомним, что каждый новый посетитель, которого предстоит идентифицировать, был представлен тремя образами, а именно 3а, 3б и 3с. Процедура их классификации выполняется последовательно: первым обрабатывается образ 3а, последним — образ 3с. Поэтому, если в процессе классификации два любых образа (например, 3а и 3б, 3а и 3с или 3б и 3с) будут отнесены к одному и тому же классу, можно считать, что идентификация прошла успешно. При этом если к одному и тому же классу отнесены первые два образа (а именно, 3а и 3б), то процесс дальнейшей идентификации можно прервать.

Синтез каждой системы «Visitor Identification» заканчивается этапом идентификации контрольных образов из ВБ. Именно на этом этапе проверяется правильность функционирования системы. Идентифицируемые контрольные образы естественно должны быть отнесены к своим классам. При этом дополнительным критерием правильности функционирования системы могут служить численные значения оценок  $\{Ko(i), Kp(j), So(l), Sp(m), Fo(h), Fp(g)\}$ .

На рис. 6.3.4 представлены результаты такой проверки для контрольных образов 5-го класса данных из ВБ. Нумерация рисунков соответствует обозначениям, используемым выше: образы с литерами от «а» до «с» соответствуют приближению лица к видеокамере.

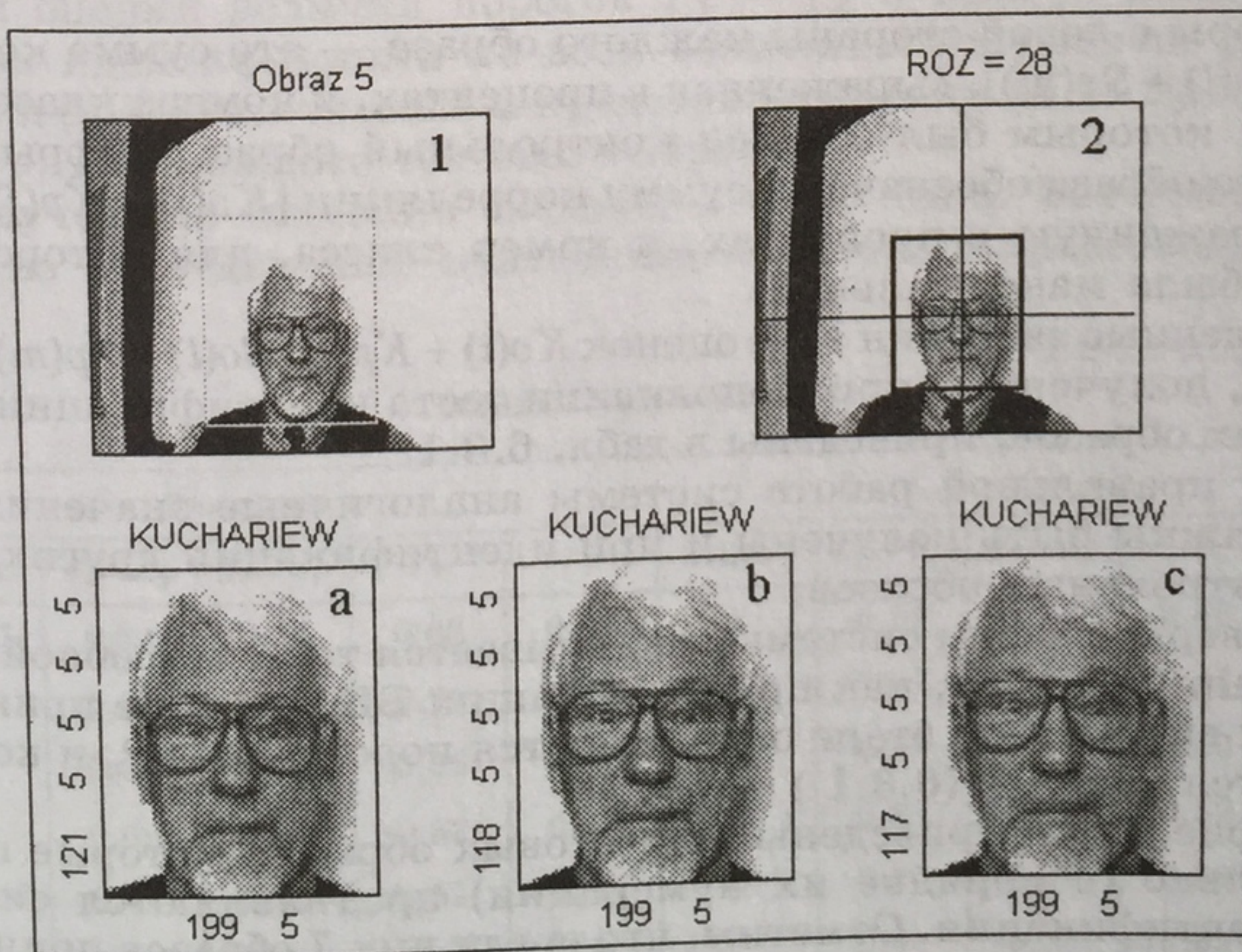


Рис. 6.3.4. Результат идентификации контрольных образов 5-го класса



Т а б л и ц а 6.3.1

Результаты идентификации контрольных образов

| Для образа 3a |           |           |      |      | Для образа 3b |           |           |      |      |
|---------------|-----------|-----------|------|------|---------------|-----------|-----------|------|------|
| КЛАСС         | $Ko + Kp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Ko + Kp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1             | 0,55      | 0,43      | 0,58 | 0,32 | 1             | 0,54      | 0,43      | 0,58 | 0,28 |
| 2             | 0,46      | 0,39      | 0,54 | 0,24 | 2             | 0,49      | 0,40      | 0,53 | 0,19 |
| 3             | 0,49      | 0,41      | 0,63 | 0,35 | 3             | 0,54      | 0,41      | 0,64 | 0,32 |
| 4             | 0,47      | 0,39      | 0,59 | 0,18 | 4             | 0,53      | 0,43      | 0,60 | 0,15 |
| 5             | 1,99      | 1,21      | 1,00 | 1,00 | 5             | 1,99      | 1,18      | 1,00 | 1,00 |
| 6             | 0,44      | 0,37      | 0,88 | 0,58 | 6             | 0,41      | 0,37      | 0,59 | 0,23 |
| 7             | 0,50      | 0,38      | 0,88 | 0,50 | 7             | 0,46      | 0,38      | 0,57 | 0,19 |

| Для образа 3c |           |           |      |      |
|---------------|-----------|-----------|------|------|
| КЛАСС         | $Ko + Kp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1             | 0,49      | 0,42      | 0,60 | 0,29 |
| 2             | 0,48      | 0,39      | 0,55 | 0,21 |
| 3             | 0,53      | 0,41      | 0,63 | 0,33 |
| 4             | 0,49      | 0,41      | 0,59 | 0,16 |
| 5             | 1,99      | 1,17      | 1,00 | 1,00 |
| 6             | 0,48      | 0,39      | 0,60 | 0,24 |
| 7             | 0,50      | 0,39      | 0,59 | 0,18 |

Цифры с левой стороны каждого образа — это сумма корреляции  $[So(l) + Sp(m)]$ , выраженная в процентах, и номера классов  $\{i, j, h, g\}$ , к которым был отнесен контрольный образ. Цифры внизу каждого образа обозначают сумму корреляции  $[Ko(i) + Kp(j)]$ , также выраженную в процентах, и номер класса, для которого эта сумма была максимальной.

Численные значения всех оценок  $Ko(i) + Kp(j)$ ,  $So(l) + Sp(m)$ ,  $Fo(h)$  и  $Fp(g)$ , полученные при выполнении теста идентификации контрольных образов, приведены в табл. 6.3.1.

При правильной работе системы аналогичные значения оценок должны быть получены и при идентификации других классов контрольных образов.

Проверка работы системы заканчивается тестом пробной идентификации образов, как принадлежащих ВБ, так и не принадлежащих ей. На этом этапе определяются пороги  $P_1 - P_4$  и корректируются условия (6.3.1) — (6.3.4).

На рис. 6.3.5 приведены 15 тестовых образов, которые последовательно (в порядке их нумерации) предъявляются системе для идентификации. Отметим, что среди них 7 образов принадлежат санкционированным пользователям, (1, 3, 5, 6, 8, 11 и 15). Однако только образ с номером 15 отражает исходные данные

своего к  
ния, хр  
совсем  
получен  
в разных  
Для  
условия  
(6.3.1) и  
 $Fo$  и  $Fp$   
Полу  
стоит и

| КЛАСС    |
|----------|
| 1 (1111) |
| 2        |
| 3        |
| 4        |
| 5        |
| 6        |
| 7        |





Рис. 6.3.5. 15 тестовых образов

своего класса, поскольку получен из соответствующего изображения, хранящегося в ИБ. Остальные образы (1, 3, 5, 6, 8 и 11) не совсем точно соответствуют образам в ИБ (или ВБ), поскольку получены в разное время, при разных освещении и фоне, а также в разных условиях приближения к камере.

Для оценки величин порогов  $P_1 - P_4$  и выбора конкретного условия идентификации из всех возможных вариантов условий (6.3.1) и (6.3.4) вычислим и представим оценки  $\{Kp + Sp, So + Sp, Fo$  и  $Fp\}$  для каждого тестового образа.

Полученные значения сведены в табл. 6.3.2. Вся таблица состоит из 15 отдельных подтаблиц, в которых представлены ре-

Таблица 6.3.2

Оценка мер подобия

| Для образа 1 |           |           |      |      | Для образа 2 |           |           |      |      |
|--------------|-----------|-----------|------|------|--------------|-----------|-----------|------|------|
| КЛАСС        | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС        | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1 (1111)     | 0,84      | 0,76      | 0,68 | 0,45 | 1            | 0,39      | 0,39      | 0,46 | 0,25 |
| 2            | 0,55      | 0,54      | 0,71 | 0,34 | 2            | 0,41      | 0,36      | 0,45 | 0,27 |
| 3            | 0,63      | 0,54      | 0,55 | 0,36 | 3            | 0,39      | 0,36      | 0,56 | 0,26 |
| 4            | 0,60      | 0,49      | 0,52 | 0,28 | 4            | 0,38      | 0,37      | 0,52 | 0,35 |
| 5            | 0,42      | 0,39      | 0,51 | 0,40 | 5            | 0,38      | 0,36      | 0,59 | 0,11 |
| 6            | 0,51      | 0,56      | 0,63 | 0,41 | 6            | 0,44      | 0,37      | 0,53 | 0,25 |
| 7            | 0,64      | 0,55      | 0,45 | 0,32 | 7            | 0,43      | 0,39      | 0,51 | 0,32 |



Продолжение табл. 6.3.2

| Для образа 3 |           |           |      |      | Для образа 4  |           |           |      |      |
|--------------|-----------|-----------|------|------|---------------|-----------|-----------|------|------|
| КЛАСС        | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1            | 0,72      | 0,58      | 0,55 | 0,40 | 1             | 0,66      | 0,66      | 0,65 | 0,33 |
| 2            | 0,79      | 0,59      | 0,57 | 0,56 | 2             | 0,60      | 0,56      | 0,65 | 0,49 |
| 3 (3313)     | 1,22      | 0,71      | 0,57 | 0,72 | 3             | 0,78      | 0,54      | 0,63 | 0,37 |
| 4            | 0,67      | 0,47      | 0,55 | 0,49 | 4             | 0,58      | 0,46      | 0,64 | 0,49 |
| 5            | 0,46      | 0,38      | 0,56 | 0,21 | 5             | 0,43      | 0,40      | 0,58 | 0,19 |
| 6            | 0,71      | 0,51      | 0,62 | 0,53 | 6             | 0,71      | 0,64      | 0,69 | 0,37 |
| 7            | 0,66      | 0,50      | 0,48 | 0,48 | 7             | 0,79      | 0,67      | 0,58 | 0,45 |
| Для образа 5 |           |           |      |      | Для образа 6  |           |           |      |      |
| КЛАСС        | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1            | 0,44      | 0,41      | 0,57 | 0,23 | 1             | 0,61      | 0,57      | 0,68 | 0,39 |
| 2            | 0,43      | 0,39      | 0,52 | 0,21 | 2             | 0,94      | 0,69      | 0,73 | 0,47 |
| 3            | 0,41      | 0,38      | 0,62 | 0,25 | 3             | 0,65      | 0,56      | 0,66 | 0,55 |
| 4            | 0,50      | 0,43      | 0,57 | 0,18 | 4             | 0,62      | 0,49      | 0,68 | 0,42 |
| 5 (5555)     | 1,08      | 0,84      | 0,81 | 0,53 | 5             | 0,42      | 0,38      | 0,57 | 0,27 |
| 6            | 0,42      | 0,38      | 0,54 | 0,29 | 6 (6666)      | 1,11      | 0,75      | 0,74 | 0,76 |
| 7            | 0,51      | 0,40      | 0,54 | 0,22 | 7             | 0,69      | 0,57      | 0,65 | 0,45 |
| Для образа 7 |           |           |      |      | Для образа 8  |           |           |      |      |
| КЛАСС        | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1            | 0,51      | 0,51      | 0,62 | 0,35 | 1             | 0,63      | 0,58      | 0,55 | 0,31 |
| 2            | 0,45      | 0,49      | 0,59 | 0,35 | 2             | 0,59      | 0,63      | 0,55 | 0,45 |
| 3            | 0,58      | 0,45      | 0,70 | 0,45 | 3             | 0,51      | 0,53      | 0,71 | 0,34 |
| 4            | 0,57      | 0,48      | 0,62 | 0,32 | 4             | 0,63      | 0,59      | 0,79 | 0,44 |
| 5            | 0,37      | 0,38      | 0,71 | 0,30 | 5             | 0,48      | 0,43      | 0,57 | 0,18 |
| 6            | 0,51      | 0,48      | 0,61 | 0,46 | 6             | 0,63      | 0,64      | 0,76 | 0,39 |
| 7            | 0,44      | 0,42      | 0,61 | 0,36 | 7 (7777)      | 1,09      | 0,87      | 0,86 | 0,51 |
| Для образа 9 |           |           |      |      | Для образа 10 |           |           |      |      |
| КЛАСС        | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1            | 0,71      | 0,57      | 0,59 | 0,52 | 1             | 0,54      | 0,55      | 0,58 | 0,34 |
| 2            | 0,97      | 0,66      | 0,61 | 0,50 | 2             | 0,64      | 0,53      | 0,60 | 0,35 |
| 3            | 0,74      | 0,54      | 0,53 | 0,54 | 3             | 0,62      | 0,52      | 0,65 | 0,43 |
| 4            | 0,80      | 0,50      | 0,61 | 0,41 | 4             | 0,52      | 0,47      | 0,62 | 0,37 |
| 5            | 0,37      | 0,35      | 0,49 | 0,30 | 5             | 0,44      | 0,38      | 0,66 | 0,32 |
| 6            | 0,86      | 0,67      | 0,63 | 0,57 | 6             | 0,47      | 0,49      | 0,64 | 0,46 |
| 7            | 0,83      | 0,61      | 0,57 | 0,38 | 7             | 0,67      | 0,55      | 0,54 | 0,38 |

результаты  
 из ВБ.  
 Каждая  
 рых разм  
 для соотв  
 мальные з  
 делены по  
 речисленн  
 мумы оце  
 Привед  
 вать велич



Продолжение табл. 6.3.2

| Для образа 11 |           |           |      |      | Для образа 12 |           |           |      |      |
|---------------|-----------|-----------|------|------|---------------|-----------|-----------|------|------|
| КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1             | 0,52      | 0,65      | 0,82 | 0,45 | 1             | 0,36      | 0,36      | 0,34 | 0,14 |
| 2 (2222)      | 0,90      | 0,85      | 0,87 | 0,64 | 2             | 0,39      | 0,36      | 0,34 | 0,16 |
| 3             | 0,67      | 0,53      | 0,75 | 0,56 | 3             | 0,40      | 0,37      | 0,46 | 0,13 |
| 4             | 0,85      | 0,64      | 0,78 | 0,47 | 4             | 0,38      | 0,35      | 0,42 | 0,15 |
| 5             | 0,50      | 0,45      | 0,65 | 0,25 | 5             | 0,38      | 0,35      | 0,47 | 0,11 |
| 6             | 0,70      | 0,70      | 0,82 | 0,57 | 6             | 0,38      | 0,37      | 0,38 | 0,10 |
| 7             | 0,72      | 0,66      | 0,74 | 0,45 | 7             | 0,37      | 0,35      | 0,47 | 0,15 |
| Для образа 13 |           |           |      |      | Для образа 14 |           |           |      |      |
| КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ | КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |
| 1             | 0,37      | 0,38      | 0,52 | 0,23 | 1             | 0,44      | 0,47      | 0,58 | 0,38 |
| 2             | 0,40      | 0,36      | 0,51 | 0,25 | 2             | 0,51      | 0,44      | 0,59 | 0,44 |
| 3             | 0,40      | 0,36      | 0,58 | 0,23 | 3             | 0,49      | 0,48      | 0,59 | 0,43 |
| 4             | 0,45      | 0,39      | 0,51 | 0,24 | 4             | 0,57      | 0,45      | 0,63 | 0,39 |
| 5             | 0,38      | 0,41      | 0,68 | 0,25 | 5             | 0,43      | 0,44      | 0,61 | 0,24 |
| 6             | 0,43      | 0,39      | 0,55 | 0,25 | 6             | 0,73      | 0,57      | 0,58 | 0,45 |
| 7             | 0,45      | 0,39      | 0,55 | 0,22 | 7             | 0,52      | 0,50      | 0,57 | 0,45 |
| Для образа 15 |           |           |      |      |               |           |           |      |      |
| КЛАСС         | $Kp + Sp$ | $So + Sp$ | $Fo$ | $Fp$ |               |           |           |      |      |
| 1             | 0,63      | 0,50      | 0,54 | 0,41 |               |           |           |      |      |
| 2             | 0,76      | 0,58      | 0,54 | 0,53 |               |           |           |      |      |
| 3             | 0,72      | 0,58      | 0,64 | 0,50 |               |           |           |      |      |
| 4 (4474)      | 1,16      | 0,79      | 0,77 | 0,82 |               |           |           |      |      |
| 5             | 0,44      | 0,37      | 0,57 | 0,19 |               |           |           |      |      |
| 6             | 0,84      | 0,62      | 0,66 | 0,53 |               |           |           |      |      |
| 7             | 0,74      | 0,66      | 0,70 | 0,62 |               |           |           |      |      |

результаты оценки мер подобия 15 тестовых образов с 7 образами из ВБ.

Каждая выделенная подтаблица содержит по 7 строк, в которых размещены вычисленные оценки  $\{Kp + Sp, So + Sp, Fo$  и  $Fp\}$  для соответствующих классов. Результаты, определяющие максимальные значения вычисленных оценок для каждого образа, выделены полужирным шрифтом. В выделенных строках также перечислены номера классов  $\{i, j, h, q\}$ , которым принадлежат максимумы оценок для конкретного тестового образа.

Приведенные в табл. 6.3.2 оценки позволяют теперь обосновать величины порогов  $P_1 - P_4$ . Фактически эти величины содер-





Рис. 6.3.6. Тестовые образы, успешно прошедшие идентификацию (верхний ряд), и образы-эталоны, хранящиеся в ВБ (нижний ряд)

жаты в первых двух столбцах таблиц  $(Kp + Sp)$  и  $(So + Sp)$ . Итак, пусть  $P_1 = 1,0$ ;  $P_2 = 0,7$ ;  $P_3 = 0,65$ ;  $P_4 = 0,6$ . Дальнейший анализ данных этой таблицы показывает, что успешная идентификация тестовых образов возможна при использовании любого из условий (6.3.1) – (6.3.4). Воспользуемся в таком случае условиями (6.3.2) и (6.3.3).

Результат идентификации тестовых образов при заданных порогах и выбранных условиях (который прежде всего необходимо рассматривать как результат тестирования системы) представлен на рис. 6.3.6 и 6.3.7.

При тестировании системы была одновременно поставлена задача отсортировать по номерам классов те образы, которые успешно прошли идентификацию. Именно этот результат представляет рис. 6.3.6.

Рассмотрим еще один пример — распознавание близнецов. В исходную базу были включены изображения двух братьев-близнецов (только по одному изображению на каждого). Из десяти предъявленных для идентификации образов братьев (по пять для каждого), семь были идентифицированы правильно. Два из трех

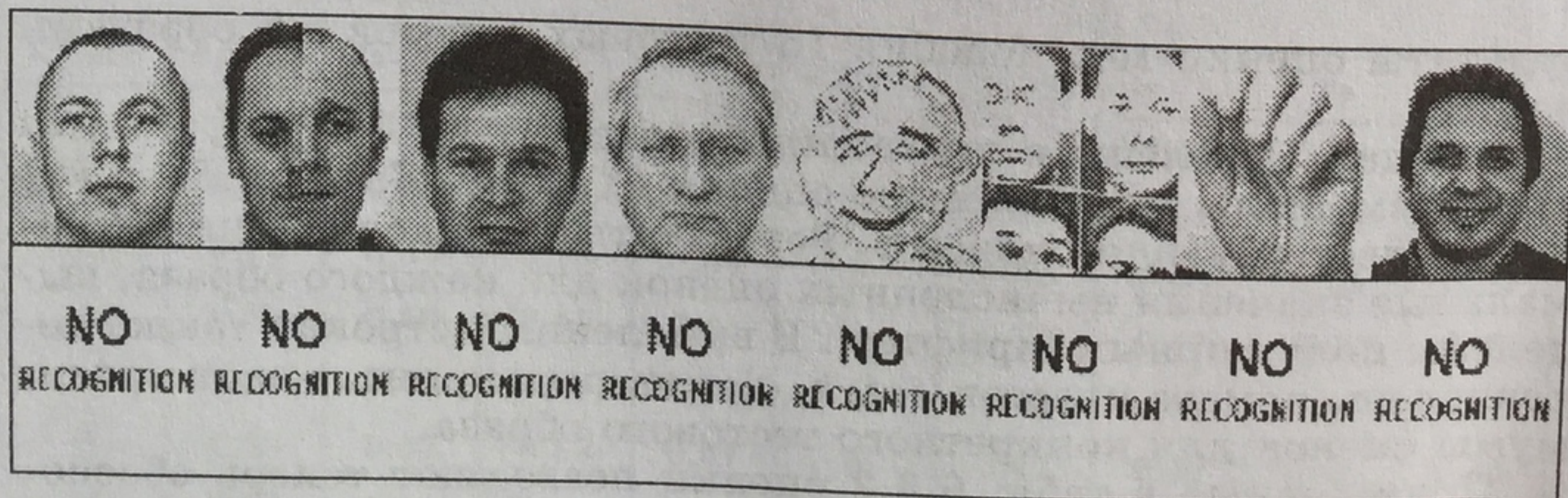
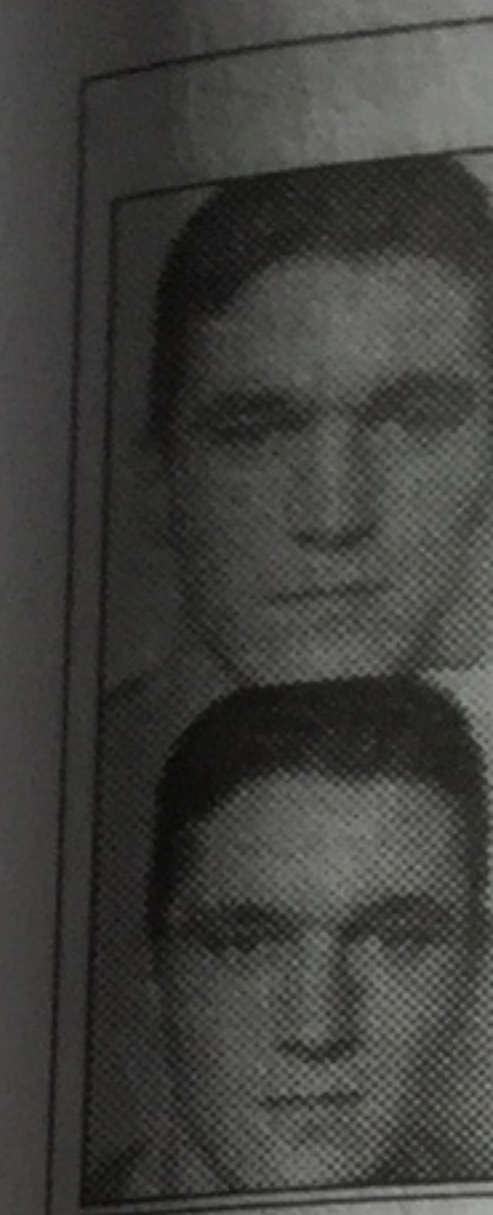


Рис. 6.3.7. Тестовые образы, не прошедшие идентификацию (верхний ряд)



оставших  
фицирова  
рис. 6.3.8  
познаван

А тепе  
[Orl\*\*] и  
рассматри  
при выбра  
вариантах  
образов не  
образов В

Замети  
два образа  
как прина  
ная акцеп  
дует из ана  
порога  $P_4$

Как вид  
Identificati  
ственной ба  
ном конкре

В то же  
FRR также  
жащих к б  
если образ  
образу в ба  
будет актуа  
ошибочный  
ется образ,  
ях, определ  
что и соотв  
условий:

- неизмен
- одинако



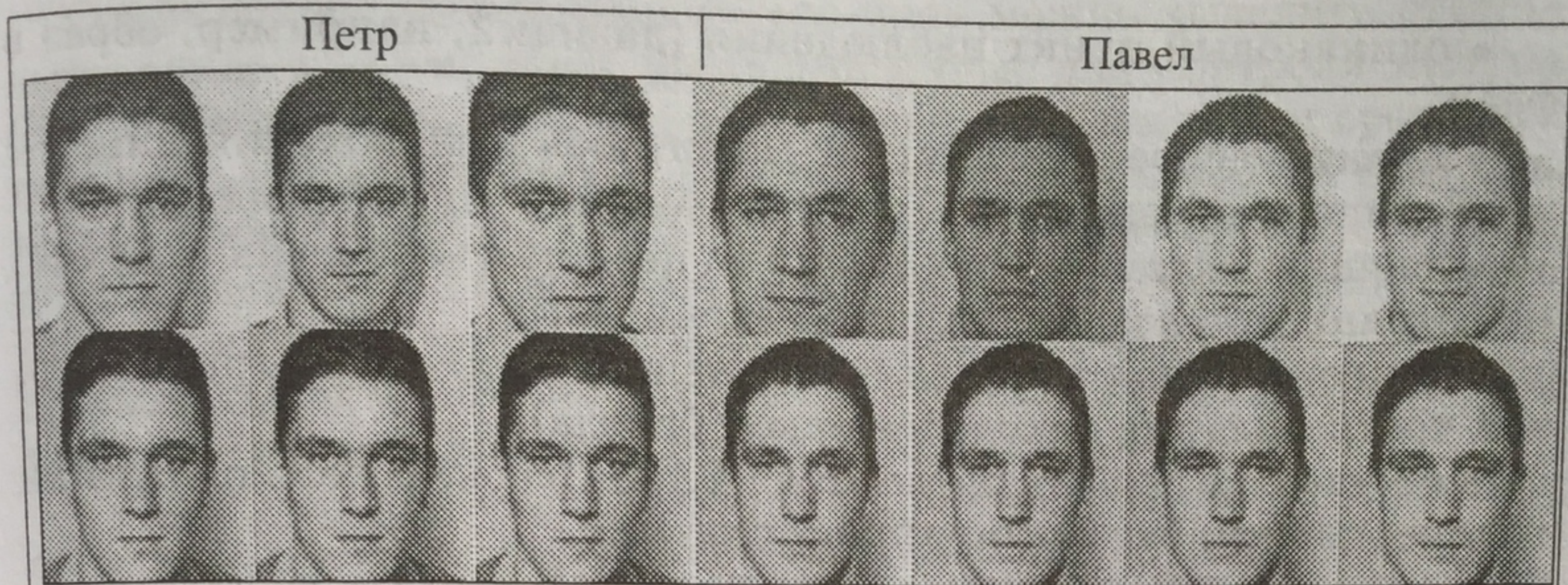


Рис. 6.3.8. Результат идентификации братьев-близнецов

оставшихся не были идентифицированы вообще, а один идентифицирован неверно. Результаты эксперимента приведены на рис. 6.3.8. Здесь верхний ряд — образы, предъявленные для распознавания, а нижний ряд — образы из базы.

А теперь в качестве тестовых образов возьмем образы из базы [Orl\*\*] и проверим, насколько они «похожи» на образы из ВБ рассматриваемой системы «Visitor Identification». Оказалось, что при выбранных ранее порогах  $P_2 = 0,7$ ,  $P_3 = 0,65$  и  $P_4 = 0,6$  и трех вариантах условий (6.3.2) – (6.3.4) ни один из предъявленных образов не был идентифицирован (не найдено подобия их среди образов ВБ) как пользователь системы.

Заметим, что при снижении порога  $P_4$  до величины 0,5 только два образа из 20 классов образов базы [Orl\*\*] были «распознаны» как принадлежащие к пользователям системы. При этом ложная акцептация произошла по условию (6.3.4). Однако, как следует из анализа данных в приведенных выше таблицах, значение порога  $P_4 = 0,5$  не может быть использовано на практике.

Как видно из приведенных примеров, новая система «Visitor Identification» не идентифицирует образы, не принадлежащие собственной базе. Из этого факта следует, что показатель FAR в данном конкретном случае стремится к нулю.

В то же время нельзя определенно утверждать, что показатель FRR также стремится к нулю. Распознавание образов, принадлежащих к базовым, поддерживается на уровне близком к 100 %, если образ «посетителя» является актуальным по отношению к образу в базе. Если образ, предъявленный для распознавания, не будет актуальным по отношению к образу в ИБ, может наступить ошибочный отказ. При этом под актуальным здесь подразумевается образ, который получен при некоторых одинаковых условиях, определяющих его (образа) биометрические характеристики, что и соответствующий образ ИБ. Отметим некоторые из этих условий:

- неизменные условия освещенности перед камерой;
- одинаковый окружающий фон вокруг изображения лица;



- одинаковый пункт наблюдения (дающий, например, образ в фас);
- практическое соответствие масштабов изображений лиц;
- одинаковое эмоциональное выражение лица (мимика);
- внешнее сходство двух лиц (или несущественные их отличия, не связанные с возрастными изменениями);
- относительная неизменность внешнего вида лица (например, той части лиц, где находятся усы и борода);
- отсутствие разных предметов на лице (серьги, кольца, очки) и голове (шапки, банты, шляпки и платки);
- относительная неизменность причесок и цвета волос;
- и т. д.

Все перечисленные условия получения исходных данных можно разделить на две группы: первая определяет как бы внешние, а вторая — внутренние причины изменчивости биометрических характеристик образов. Заметим, что все отмеченное усугубляется тем, что ИБ данных содержит только по одному образу в каждом классе. При наличии 3 – 5 образов в каждом классе ИБ требования к актуальности образов могут быть существенно смягчены, что также повлечет за собой снижение показателя FRR до значений, приближающихся к нулю.

Наконец, использование PIN, играющего роль последней инстанции на этапе верификации результата распознавания, позволит довести уровень идентификации «посетителя» до 100 %, а FRR до нуля. При этом в качестве самого простого PIN может быть использован закодированный номер класса, к которому принадлежит санкционированный пользователь, который известен только ему самому. Запрос на подачу PIN система выдает либо безусловно при каждом сеансе, либо в случае неоднозначности результата распознавания и невозможности принятия окончательного решения.

#### 6.4. СИСТЕМЫ, ОБЕСПЕЧИВАЮЩИЕ КОНТРОЛЬ ДОСТУПА

В описанной выше системе «Visitor Identification» в каждом классе ИБ имелся только один образ, а в качестве исходных признаков использовались яркостные признаки. Собственно распознавание выполнялось в пространстве исходных признаков, сильно зависящих от условий получения образов. Именно поэтому распознавание санкционированного посетителя выполняется успешно, если его образ является актуальным по отношению к образу в ИБ данных. При этом под актуальным здесь подразумевается образ, который имеет с соответствующим образом ИБ одинаковые в некотором смысле характеристики. Для рассмотренного класса систем это требование является обязательным, в противном случае происходит ошибочный отказ.



Среди систем распознавания «on-line» можно выделить отдельный подкласс — системы, предназначенные для контроля доступа (КД) к чему-нибудь: например, к компьютеру, семейным средствам передвижения (например, к автомобилю, яхте, самолету), базе данных, индивидуальным и коллективным счетам в банках, собственному оружию или в специальные помещения. Особенности этого класса систем — длительный срок их эксплуатации и существенно повышенные (в сравнении с системами класса «Visitor Identification») требования к надежности идентификации пользователя.

Очевидно, что в приложении к системам КД неактуальность образа пользователя становится главной проблемой. Последнее объясняется тем, что при длительной эксплуатации системы КД практически невозможно выдержать одни и те же условия получения исходных данных.

Очевидно и то, что полное решение возникших при этом проблем невозможно, поскольку априори невозможно предусмотреть все случаи вариантов изменений исходных данных. Однако можно попробовать построить такую систему распознавания, которая будет поприбаву по отношению к некоторым изменениям условий получения исходных данных и некоторым изменениям биометрических характеристик образов — пользователей систем. Такая система распознавания может быть реализована на подходах, описанных в гл. 5 и разд. 6.2 и 6.3.

Суть этих подходов состоит в следующем.

В качестве исходных признаков выберем градиенты яркости, что позволит существенно элиминировать влияние изменений условий освещенности образов в момент сеанса с видеокамерой. И хотя теперь индивидуальные черты лиц представляются через пространство новых признаков — градиентов яркости, в данном случае все еще сильно ощущается непосредственное влияние других внешних и внутренних условий получения исходных образов на их биометрические характеристики.

Поэтому для уменьшения этого непосредственного влияния перейдем от пространства исходных признаков к редуцированному пространству признаков. В нем и будем выполнять распознавание.

Для того чтобы элиминировать влияние разности масштабов распознаваемых образов и образов-эталонов на результат распознавания, сформируем в ВБ (и в дальнейшем будем использовать) несколько образов-эталонов, имитирующих приближение/удаление к видеокамере в момент сеанса с ней.

И, наконец, в каждом из образов будем выбирать несколько подобразов, определяющих наиболее стабильную часть лица человека, чтобы уменьшить влияние внутренних причин изменчивости на биометрические характеристики распознаваемого образа.

Теперь рассмотрим конкретный пример системы КД, предназначенной для семи пользователей (семь пользователей было выбрано для того, чтобы можно было сравнить эту систему с рассмотренной ранее системой «Visitor Identification»).



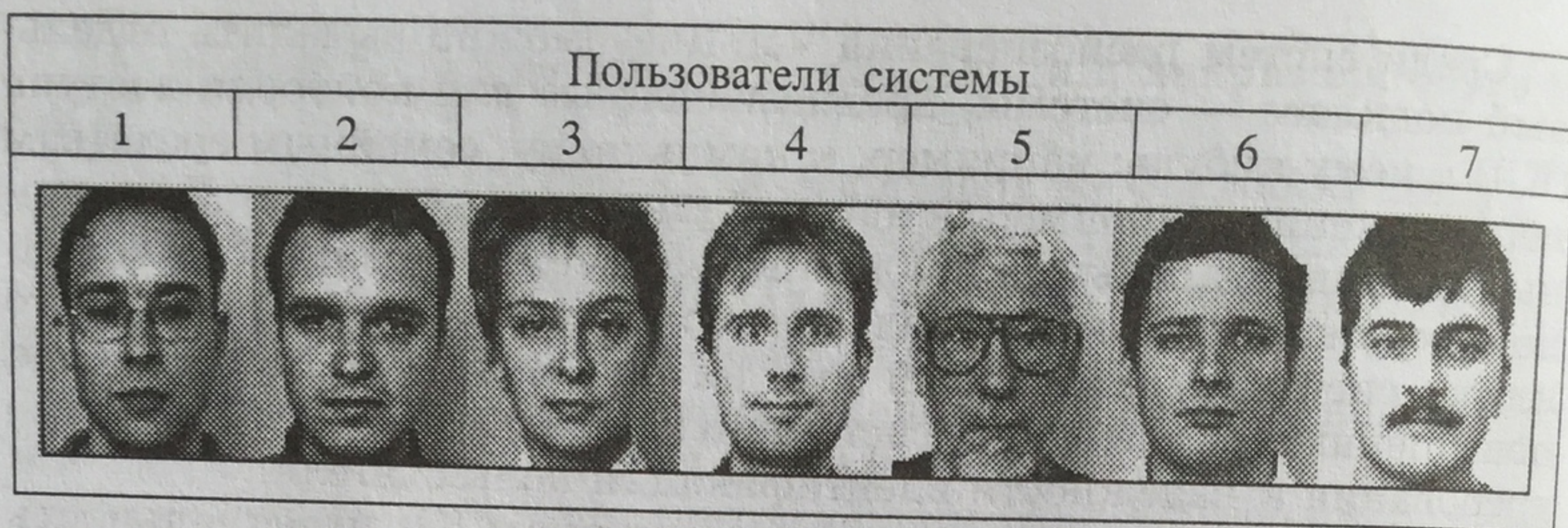


Рис. 6.4.1. Санкционированные пользователи системы

Итак, исходная база данных содержит 7 классов образов и в каждом классе имеет по одному образу. Все семь санкционированных пользователей системы представлены на рис. 6.4.1.

Процесс синтеза образов ВБ на примере исходных данных первого класса представлен на рис. 6.4.2.

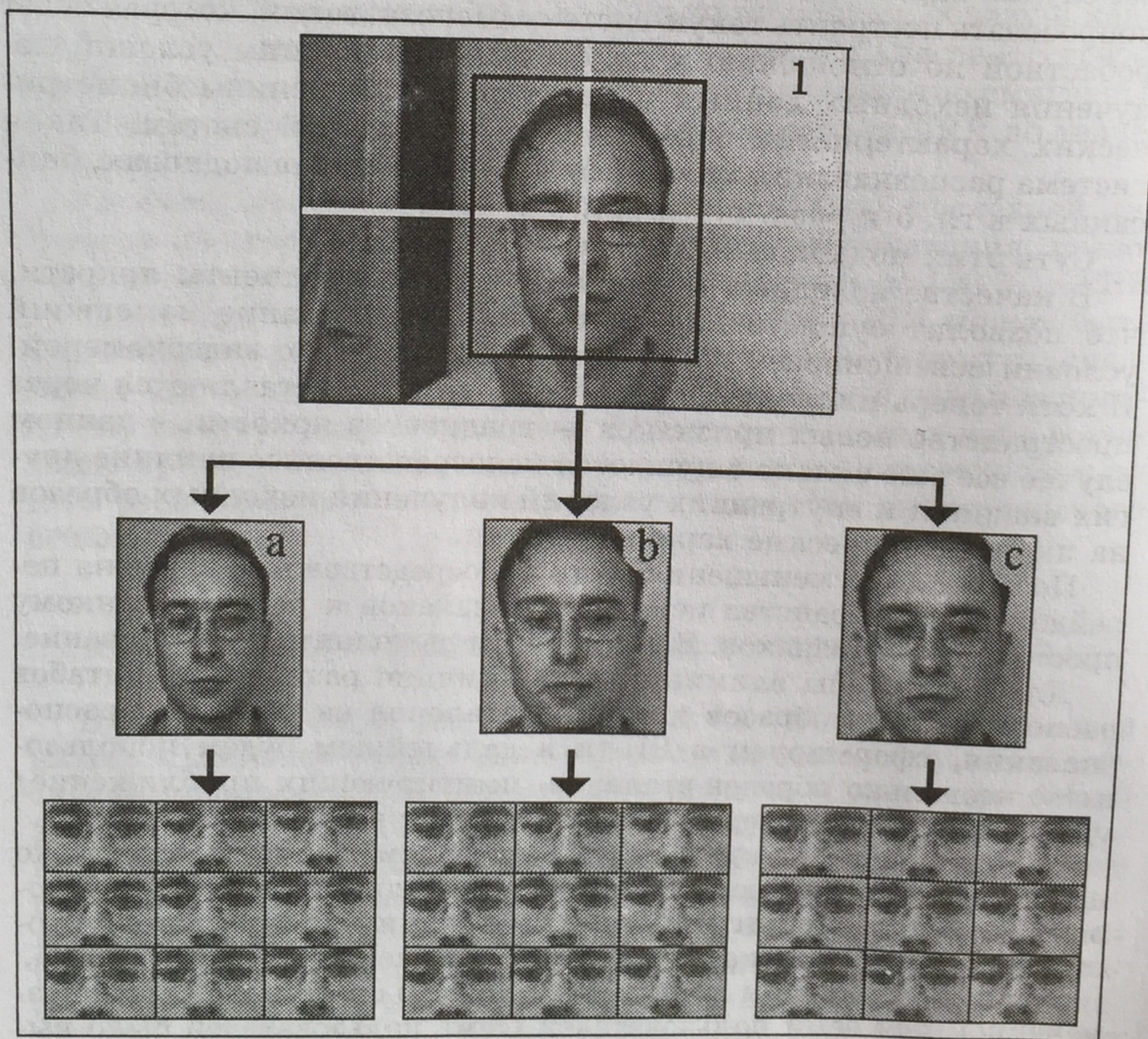


Рис. 6.4.2. Процесс синтеза данных ВБ на примере исходного образа первого класса



Здесь из единственного изображения, полученного непосредственно с видеокамеры, выделяется первоначальный образ лица. На этом образе определяются координаты центров глаз и расстояние  $R$  между ними. Далее на основе этих оценок определяются ось симметрии лица и положение линии глаз на исходном изображении (помечено цифрой 1).

Затем относительно этой оси симметрии и положения линии глаз определяется базовая рамка размером  $112 \times 96$  пикселей, в поле которой находится образ лица. Выбирая в соответствии с алгоритмом, представленным в разд. 6.2.2, еще два подобных образа с рамками меньшего и большего (относительно базового) размера и перемасштабируя их к одному и тому же размеру ( $112 \times 96$  пикселей), можно получить 3 образа лица (образы 3a, 3b и 3c) с разными значениями величины  $R$ . Этот ряд образов имитирует приближение и/или удаление посетителя к камере в момент сеанса с ним.

На каждом из образов 3a, 3b и 3c выбирается исходное положение некоторой рамки, определяющей размер выбираемого с ее помощью подобраза. Далее выполняется сканирование с помощью этой рамки данных образов по значениям координат, которые изменяются относительно оси симметрии лица и линии глаз (см. в разд. 6.2). При этом из каждого образа 3a, 3b и 3c получаем по 9 подобразов. Описанный процесс представлен схематически в нижней части рис. 6.4.2.

Естественно, что синтез данных ВБ повторяется для всех  $K$  классов образов. В рассматриваемом нами случае он повторялся бы 7 раз.

Теперь переходим к градиентам яркости. Для каждого из подобразов вычисляются три градиента (рис. 6.4.3).

При этом образ 1 соответствует суммарному градиенту яркости по оси  $X$  и  $Y$ , образ 2 — является градиентом яркости по оси  $X$ , а образ 3 — градиент яркости по оси  $Y$ . Здесь представлены градиенты, нормированные по (6.2.2). Признаками теперь являются значения градиентов в каждом пикселе.

Редукция исходного пространства признаков реализуется методами PCA и LDA в строгом соответствии с методикой, описанной в гл. 5. При этом от каждой матрицы градиентов остается только по  $(K - 1)$  признаку, где  $K$  — число классов образов в ИБ. Если каждая матрица градиентов имеет размер  $M \times N$  пикселей, то исходное пространство признаков уменьшается при этом в  $MN/(K - 1)$  раза.

В общем случае, по окончании процедуры синтеза новых образов и редукции пространства признаков в памяти системы содержатся следующие данные.

1. Параметры системы  $\{K, L, Q, F, G, p\}$ , где  $K$  — число классов;  $L$  — число образов-эталонов в каждом классе;  $Q$  — общее число образов в классе, включая контрольные образы;  $F$  — число подобразов для каждого образа-эталона;  $G$  — число матриц-градиентов;  $p$  — параметр метода PCA (см. разд. 5.1).
2. Значения размеров всех используемых масок.



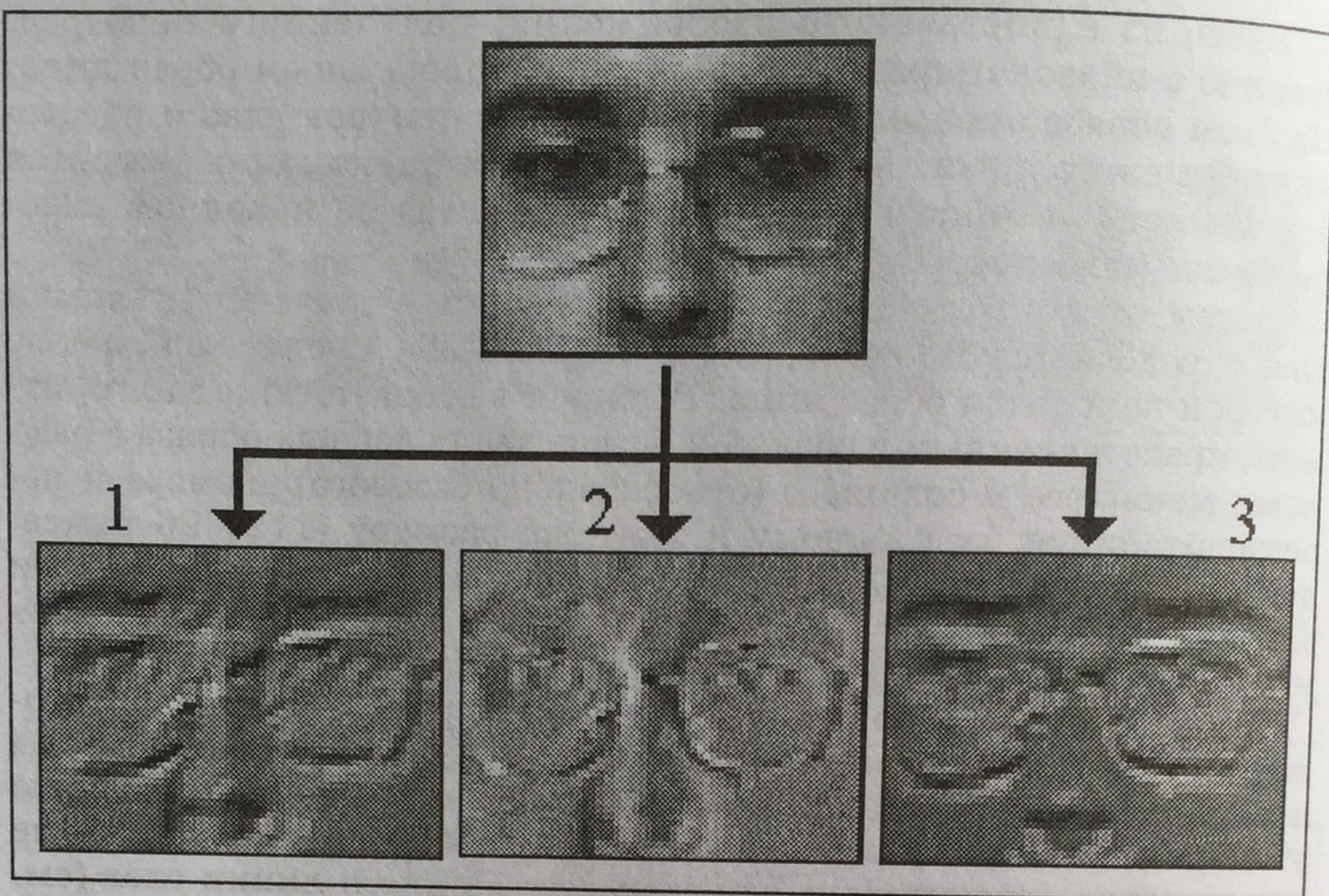


Рис. 6.4.3. Подобраз и три вычисленных для него градиента

3.  $K$  исходных изображений, полученных непосредственно с видеокамер.

4.  $K$  изображений лиц пользователей (образов-эталонов  $3b$  размером  $112 \times 96$  пикселей), полученных из исходных изображений.

5. Трехмерный массив размером  $2 \times 3 \times K$ , предназначенный для хранения значений координат  $X$  и  $Y$ , которые определяют точку пересечения оси симметрии лица и линии глаз на образах эталонах.

6. Трехмерный массив размером  $256 \times (L \times F) \times K$ , определяющий подобразы, перемасштабированные к размеру  $16 \times 16$  пикселей и записанные векторами в результате выполнения операции конкатенации столбцов.

7. Шесть двумерных массивов, определяющих матрицы редукции признаков по методу PCA (три матрицы имеют размер  $p \times M$  и предназначены для редукции матриц-градиентов по столбцам, а три другие матрицы имеют размер  $N \times p$  и предназначены для редукции матриц-градиентов по строкам);

8. Три двумерных массива, определяющих матрицы редукции признаков по методу LDA (матрицы имеют размер  $6 \times p^2$ ).

9. Двумерный массив размером  $(3 \times (K - 1)) \times (K \times L \times F)$ , определяющий признаки редуцированного пространства для всех  $K$  классов образов, трех образов ( $3a$ ,  $3b$  и  $3c$ ) в каждом классе,  $L$  подобразов для каждого из образов  $3a$ ,  $3b$  и  $3c$  и всех  $F$  матриц градиентов.

10. Двумерный массив размером  $(3 \times (K - 1)) \times K$ , определяющий «центры классов» — средние значения признаков в каждом классе.

На эт  
выполне  
ходимо  
рованно  
рые пар

Оцен  
правил  
классов  
их собст  
ный кон  
но судит  
но, о пот  
разов в

На ри  
ченное  
(в данно  
 $X$  и  $Y$ ).  
 $L = 3$ ,  $F$   
образов,  
Первые  
3D-прос

В пра  
класса  
отмечен  
ставлен  
показыв

Анал  
знака ре  
знаков д

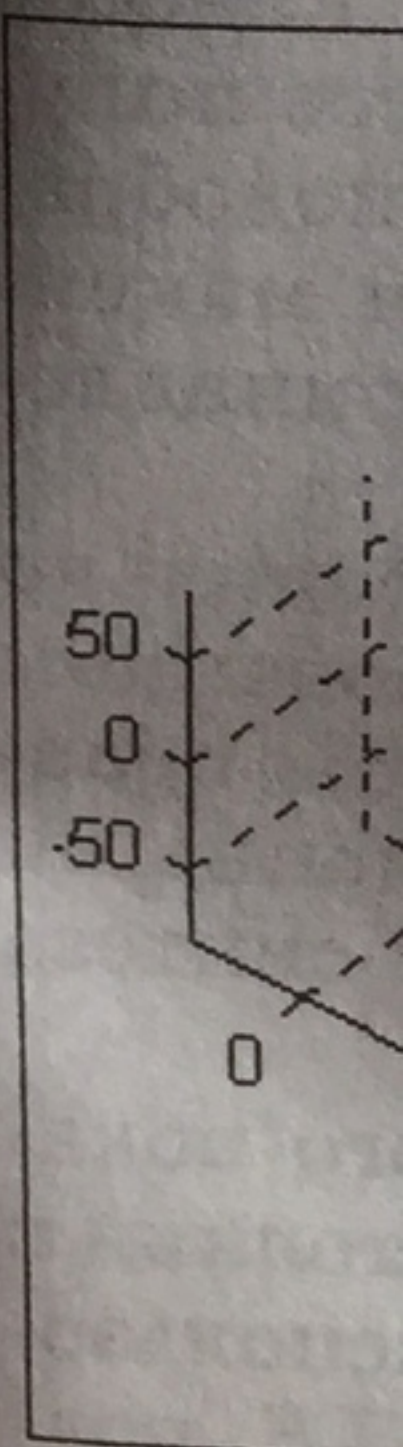


Рис. 6.4.4



На этапе синтеза системы распознавания (а конкретней, после выполнения редукции исходного пространства признаков) необходимо проверить качество кластеризации признаков в редуцированном пространстве и, при необходимости, изменить некоторые параметры системы для улучшения кластеризации [Гон78].

Оценить качество можно на основе некоторых формальных правил (например, можно оценить расстояния между центрами классов или величины средних отклонений образов в классах от их собственных центров). Однако можно использовать и визуальный контроль: по характеру расположения центров классов можно судить о качестве выполненной кластеризации и, следовательно, о потенциальной возможности собственно классификации образов в редуцированном пространстве признаков.

На рис. 6.4.4 представлено 3D-пространство признаков, полученное в результате редукции для одной из матриц градиентов (в данном случае матрицы суммарного градиента яркости по оси  $X$  и  $Y$ ). При этом параметры системы были следующими:  $K = 7$ ,  $L = 3$ ,  $F = 9$ ,  $p = 10$ . Это пространство представляет  $7 \times 3 \times 9 = 189$  образов, каждый из которых описывается всего 6 признаками. Первые три из них определяют координаты  $X$ ,  $Y$  и  $Z$  заданного 3D-пространства.

В правой стороне рис. 6.4.4 обозначены все образы первого класса (также для матрицы суммарного градиента). Цифрой 1 отмечен «центр класса», а каждый отдельный образ класса представлен точкой в 3D, соединенной с центром класса. Эта линия показывает отклонение образа от центра класса.

Аналогично можно было бы представить и следующие три признака редуцированного пространства, а также пространство признаков для других матриц градиентов.

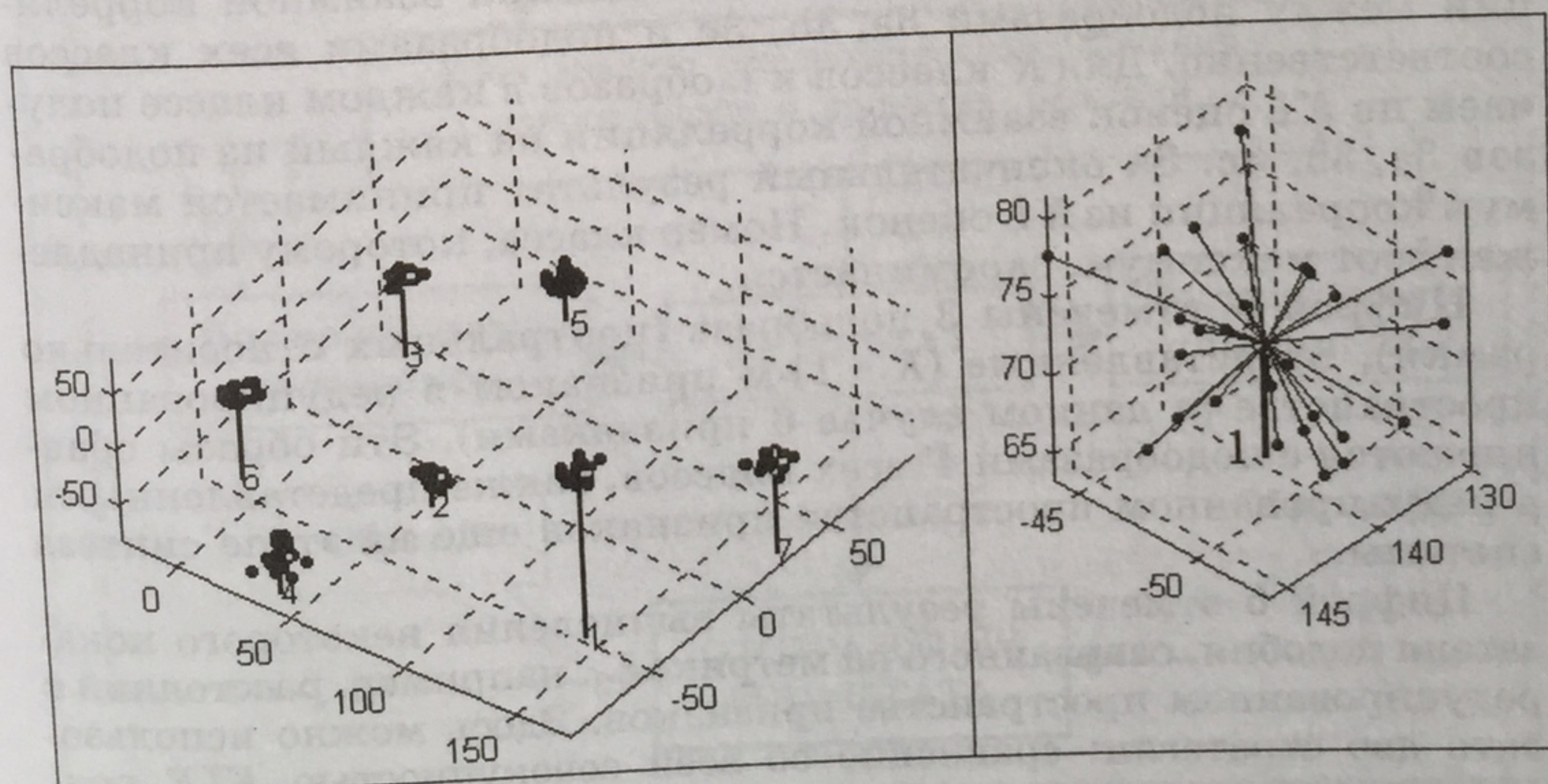


Рис. 6.4.4. Редуцированное пространство признаков, представленное в 3D-форме, для системы с параметрами  $K = 7$ ,  $L = 3$ ,  $F = 9$ ,  $p = 10$



Результат, представленный на рис. 6.4.4, можно оценить как очень хороший. В противном случае, необходимо было бы выбрать другие параметры системы, и прежде всего параметр  $p$ , что отмечалось в гл. 5.

Представим теперь структуру рассматриваемой системы (рис. 6.4.5).

Меру подобия между образами-эталоны и каждым новым образом (или контрольными образами) будем определять двумя методами: либо базирующимся на взаимной корреляции, которая будет применяться в пространстве исходных образов, либо реализованным на основе метрики  $L_2$ , которая используется в редуцированном пространстве признаков. О приоритетах этих методов скажем несколько позже.

На структурной схеме системы цифрами 1 и 2 обозначены посетитель системы КД и исходное изображение, полученное непосредственно с камеры. Цифрами 3a, 3b и 3c отмечены три выделенных из исходного изображения образа лица, предназначенных для идентификации (в режиме тестирования системы это будут контрольные образы). Все они размещаются в специально отведенном для них буфере. Образ 3b получен непосредственно из исходного изображения (без изменений его масштаба), а образы 3a и 3c на удалении от видеокамеры и при приближении лица к ней. В связи с этим, на образе 3a изображение лица будет уменьшенным, на образе 3c — увеличенным по отношению к образу 3b.

Буквами 3a, 3b, 3c отмечены перемасштабированные к размеру  $16 \times 16$  пикселей подобразы — части лиц новых (контрольных) образов 3a, 3b и 3c, представленные в форме векторов. Именно они сравниваются с подобразами 4 всех классов данных из ВБ, также представленными в форме векторов.

Цифрой 5 отмечен результат вычисления взаимной корреляции между подобразами 3a, 3b, 3c и подобразами всех классов соответственно. Для  $K$  классов и  $L$  образов в каждом классе получаем по  $KL$  оценок взаимной корреляции на каждый из подобразов 3a, 3b, 3c. За окончательный результат принимается максимум корреляции из  $KL$  оценок. Номер класса, которому принадлежал этот максимум, запоминается.

Цифрой 3' отмечены 3 подобраза (центральных относительно рамки), представленные  $(K - 1)$ -м признаком в редуцированном пространстве (в данном случае 6 признаками). Эти образы сравниваются с подобразами 4' всех классов, также представленными в редуцированном пространстве признаков еще на этапе синтеза системы.

Цифрой 6 отмечены результаты вычисления некоторого показателя подобия, основанного на метрике  $L_2$ , например, расстояний в редуцированном пространстве признаков. Здесь можно использовать две стратегии: сравнение со всей совокупностью  $KLF$  подобразов 4' или сравнение с  $K$  средними подобразами, вычисленными для совокупности 4'. В первом случае получаем по  $KLF$  оценок







показателя подобия для каждого из образов  $3'$ . За окончательный результат принимаются три оценки номера класса, которым соответствовали три минимальных расстояния из  $KLF$  оценок. Далее с помощью мажоритарного принципа (два из трех) определяется номер класса, которому принадлежит распознаваемый образ.

Во втором случае получаем по  $K$  оценок показателя подобия для каждого из образов  $3'$ . За окончательный результат принимается номер класса, которому соответствует минимум расстояния из  $K$  оценок.

Цифра 8 означает окончание процесса анализа результата классификации нового образа и принятие окончательного решения. Далее идет отображение результата в случае успешной идентификации посетителя (линия с цифрой 10), повторный запуск видеокамеры (линия с цифрой 7) в случае отрицательного результата распознавания и/или запись результата в протокол сеанса с посетителем (линия с цифрой 9).

Положим, что в данной системе основным является метод классификации, основанный на взаимной корреляции. Он достаточно хорошо себя проявил при испытаниях системы «Visitor Identification». Очевидно, что от него не стоит отказываться, и его достаточно лишь несколько (несущественно) видоизменить применительно к конкретно рассматриваемой системе.

Дополнительным методом, выполняющим функции верификации, будет в таком случае метод классификации на основе метрики  $L_2$ . Верификацию будем применять в тех случаях, когда невозможно принятие решения об окончании идентификации на основе корреляционного метода.

Если результаты классификации по основному алгоритму и верификации по дополнительному одинаковые, распознаваемый образ считается идентифицированным.

В случаях, когда результаты классификации по основному алгоритму и верификации по дополнительному являются разными, распознаваемый образ считается неидентифицированным.

Рассмотрим теперь подробнее работу блока анализа результатов классификации и принятия решений.

Будем считать, что блоком «Определение меры подобия образов на основе корреляции» вычислены коэффициенты взаимной корреляции одного подобраза группы  $3a$ ,  $3b$ ,  $3c$  со всеми подобразами группы 4 (с каждым образом и в каждом классе). Из всей совокупности полученных оценок корреляции выбираем только максимальные значения. При этом для каждого из подобразов группы  $3a$ ,  $3b$ ,  $3c$  и соответствующих им подобразов 4 получаем одну оценку максимума корреляции  $Kp$ . Всего таким образом для каждого нового (контрольного) образа получим три оценки максимума корреляции:  $Kp(i)$ ,  $Kp(j)$ ,  $Kp(h)$ , где  $i$ ,  $j$ ,  $h$  — номера классов, которым принадлежат эти максимумы.

Далее будем исходить из того, что оценки  $Kp(i)$ ,  $Kp(j)$  и  $Kp(h)$  определяют принадлежность распознаваемого образа к одному классу только в том случае, если выполняется условие



$$\max\{Kp(i), Kp(j), Kp(h)\} > P_1 \text{ при } i = j = h, \quad (6.4.1)$$

где  $P_1$  — некоторый установленный порог.

Будем также считать, что блоком «*Определение меры подобия образов на основе метрики  $L_2$* » вычислены оценки расстояний между каждым образом группы 3' со всеми образами группы 4'. При этом для каждого образа группы 3' можем получить либо  $K$  оценок (если оценка выполняется относительно центров классов), либо  $KLF$  оценок (для  $K$  классов,  $L$  образов в каждом классе и  $F$  подобразов для каждого образа). Из каждой совокупности полученных оценок определяем номер класса так, как это было описано выше. По окончании этой процедуры получим три оценки номеров классов:  $t, g, v$ , соответствующих подобразам  $a, b$  и  $c$ . Тогда, если выполняется условие  $t = g = v$ , то классификацию нового образа можно считать завершенной.

Поскольку представленная процедура является дополнительной, то окончательное принятие решения определяется при выполнении одного из условий:

$$i = j = h = t = g = v; \quad (6.4.2)$$

$$\max\{Kp(i), Kp(j), Kp(h)\} > P_2 \text{ при } (t + g + v)/3 = t; \quad (6.4.3)$$

$$\begin{aligned} &\text{любые пять индексов из } \{i, j, h, t, g, v\} \\ &\text{имеют одинаковое значение.} \end{aligned} \quad (6.4.4)$$

Рассмотрим еще одну стратегию проведения верификации. Напомним, что с каждого образа 3a (3b или 3c) можно получить по 9 подобразов, входящих в группу 3'. Будем считать, что блоком «*Определение меры подобия образов на основе метрики  $L_2$* » вычислены оценки расстояний между каждым подобразом группы 3' со всеми образами группы 4'. При этом для каждого образа группы 3' можем получить либо  $K$  оценок (если оценка выполняется относительно центров классов), либо  $KLF$  оценок (для  $K$  классов,  $L$  образов в каждом классе и  $F$  подобразов для каждого образа). Из каждой совокупности полученных оценок выберем одну, соответствующую минимуму расстояния, и запомним номер  $N_k$  класса, которому соответствует этот минимум. Далее в специальном счетчике будет накапливаться «1» так, что  $S(N_k) = S(N_k) + 1$ , где  $k = 1, 2, \dots, K$ . Принимая во внимание нулевое начальное значение всех разрядов  $S(N_k)$  счетчика перед началом процедуры оценки, сумма всех единичек в каждом счетчике будет равна величине  $F$  (в нашем случае  $F = 9$ ).

По окончании этой процедуры рассматриваем четыре оценки: номер разряда каждого из трех счетчиков (суть — номер класса), который содержит максимальное число накопленных единичек, и сумму всех единичек, соответствующих трем максимумам. Тогда образ считается правильно распознанным, если выполняются два условия:

$$t = g = v \text{ и } [S(N_t) + S(N_g) + S(N_v)] > P_2, \quad (6.4.5)$$



где  $t, g, v$  — номера разрядов счетчиков (классов по ВБ), которые содержат максимальное число накопленных единиц;  $P_2$  — некоторый установленный порог.

А теперь, заканчивая синтез системы, предъявим ей несколько образов с тем, чтобы оценить получаемые меры подобия. При этом сейчас нас интересуют только оценки на основе метрики  $L_2$ , поскольку оценки на основе корреляции можно использовать из системы «Visitor Identification». Из всех представленных выше стратегий классификации выберем последнюю, по которой можно получить довольно обширный статистический материал для оценки меры подобия.

На рис. 6.4.6 представлены образы для тестирования системы. Первые пять образов принадлежат санкционированным пользователям системы. Напомним, что в ИБ рассматриваемой системы имеется только по одному образу санкционированного пользователя. Если говорить о лице этого пользователя, то оно представлено в ИБ строго в «фас» и не имеет никаких лишних предметов (например, шапок на голове). Предъявленные образы не отвечают этому требованию. Образ номер 6 вообще является рисунком, а оставшиеся четыре образа не представляют санкционированных пользователей.

Полученные результаты представлены в табл. 6.4.1. Здесь в левой части таблицы представлены характеристики образа, например, 1(+) означает, что оценка выполнена для образа № 1 из рис. 6.4.6 и что этот образ представляет санкционированного пользователя (+), а 6(−) означает, что оценка выполнена для образа № 6 из рис. 6.4.6 и что этот образ не представляет санкционированного пользователя (−). Строки, обозначенные буквами а, b и с, определяют оценки для соответствующих подобразов. Строка «а + b + с» соответствует

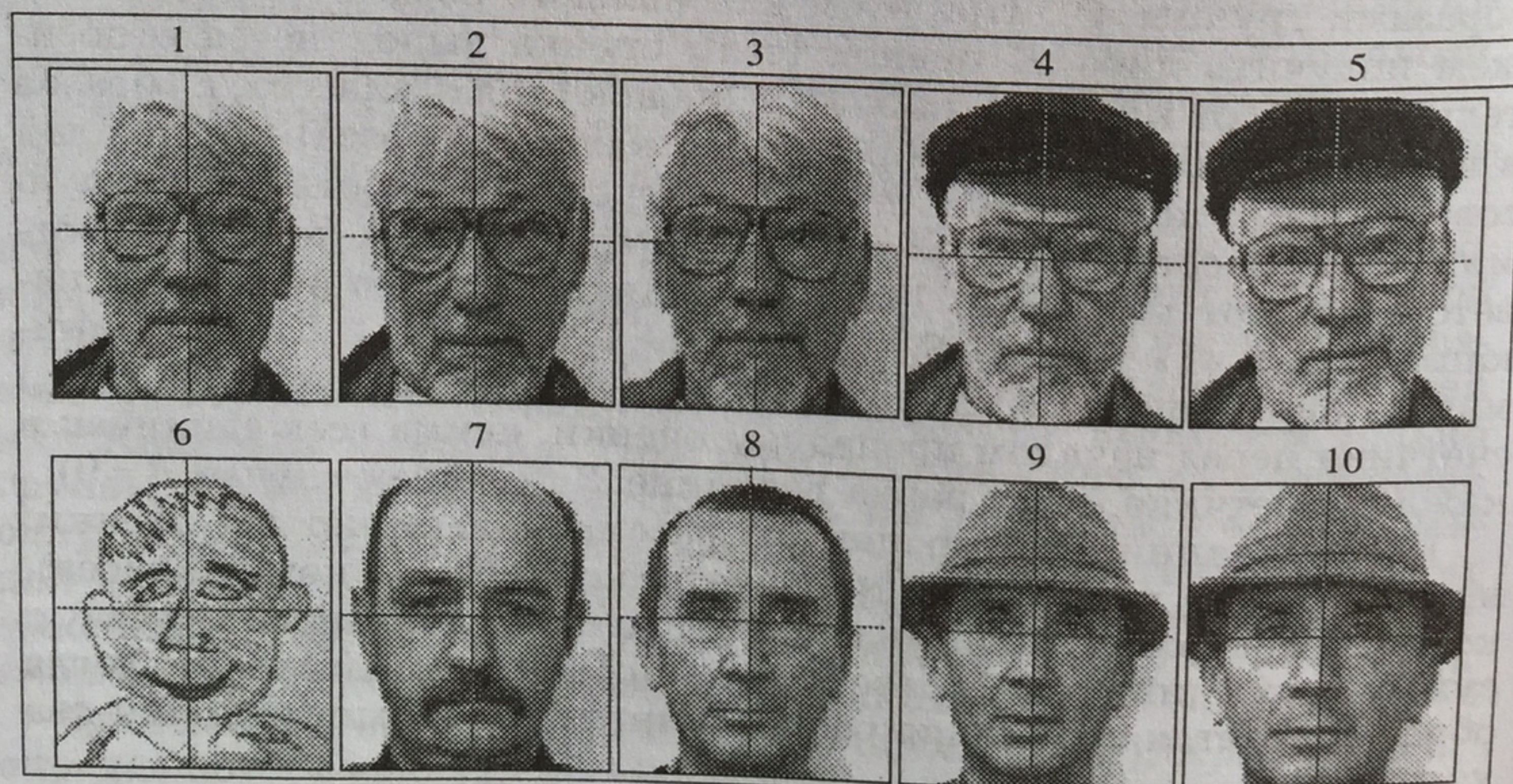


Рис. 6.4.6. Образы для тестирования системы КД



Таблица 6.4.1

## Результаты классификации тестовых образов

| Характеристика образа |           | Номер разряда (класса по ВВ) |    |   |   |    |   |   |
|-----------------------|-----------|------------------------------|----|---|---|----|---|---|
| №                     | Подобраз  | 1                            | 2  | 3 | 4 | 5  | 6 | 7 |
| 1(+)                  | a         | 0                            | 1  | 0 | 0 | 7  | 0 | 1 |
|                       | b         | 0                            | 1  | 0 | 0 | 8  | 0 | 0 |
|                       | c         | 0                            | 2  | 0 | 0 | 7  | 0 | 0 |
|                       | a + b + c | 0                            | 4  | 0 | 0 | 22 | 0 | 1 |
| 2(+)                  | a         | 0                            | 0  | 0 | 1 | 7  | 1 | 0 |
|                       | b         | 2                            | 1  | 0 | 1 | 5  | 0 | 0 |
|                       | c         | 2                            | 0  | 0 | 0 | 5  | 1 | 1 |
|                       | a + b + c | 4                            | 1  | 0 | 2 | 17 | 2 | 1 |
| 3(+)                  | a         | 0                            | 0  | 0 | 0 | 9  | 0 | 0 |
|                       | b         | 0                            | 0  | 0 | 1 | 8  | 0 | 0 |
|                       | c         | 0                            | 0  | 0 | 0 | 9  | 0 | 0 |
|                       | a + b + c | 0                            | 0  | 0 | 1 | 26 | 0 | 0 |
| 4(+)                  | a         | 0                            | 1  | 0 | 0 | 6  | 0 | 2 |
|                       | b         | 0                            | 0  | 0 | 0 | 7  | 0 | 2 |
|                       | c         | 0                            | 0  | 0 | 0 | 8  | 0 | 1 |
|                       | a + b + c | 0                            | 1  | 0 | 0 | 21 | 0 | 4 |
| 5(+)                  | a         | 0                            | 2  | 0 | 0 | 7  | 0 | 0 |
|                       | b         | 0                            | 4  | 1 | 0 | 4  | 0 | 0 |
|                       | c         | 0                            | 2  | 0 | 0 | 6  | 0 | 1 |
|                       | a + b + c | 0                            | 8  | 1 | 0 | 17 | 0 | 1 |
| 6(-)                  | a         | 2                            | 2  | 2 | 0 | 2  | 1 | 0 |
|                       | b         | 1                            | 2  | 2 | 4 | 0  | 0 | 0 |
|                       | c         | 0                            | 2  | 3 | 0 | 2  | 1 | 1 |
|                       | a + b + c | 3                            | 6  | 7 | 4 | 4  | 2 | 1 |
| 7(-)                  | a         | 0                            | 5  | 0 | 1 | 0  | 0 | 3 |
|                       | b         | 1                            | 1  | 0 | 2 | 4  | 0 | 1 |
|                       | c         | 2                            | 5  | 0 | 0 | 0  | 0 | 2 |
|                       | a + b + c | 3                            | 11 | 0 | 3 | 4  | 0 | 6 |
| 8(-)                  | a         | 0                            | 3  | 0 | 0 | 2  | 0 | 4 |
|                       | b         | 1                            | 3  | 0 | 0 | 3  | 0 | 2 |
|                       | c         | 1                            | 2  | 0 | 0 | 3  | 0 | 3 |
|                       | a + b + c | 2                            | 8  | 0 | 0 | 8  | 0 | 9 |
| 9(-)                  | a         | 1                            | 4  | 1 | 0 | 1  | 1 | 1 |
|                       | b         | 1                            | 4  | 0 | 2 | 0  | 1 | 1 |



| Характеристика образа |           | Номер разряда (класса по ВВ) |    |   |   |   |   |   |
|-----------------------|-----------|------------------------------|----|---|---|---|---|---|
| №                     | Подобраз  | 1                            | 2  | 3 | 4 | 5 | 6 | 7 |
| 9(-)                  | c         | 3                            | 0  | 0 | 0 | 1 | 0 | 5 |
|                       | a + b + c | 5                            | 8  | 1 | 2 | 2 | 2 | 7 |
| 10(-)                 | a         | 0                            | 5  | 0 | 1 | 0 | 0 | 3 |
|                       | b         | 1                            | 1  | 0 | 2 | 4 | 0 | 1 |
|                       | c         | 2                            | 5  | 0 | 0 | 0 | 0 | 2 |
|                       | a + b + c | 3                            | 11 | 0 | 3 | 4 | 0 | 6 |

суммарной оценке по трем образам. Собственно оценки представлены в правой части таблиц (для всех 7 разрядов счетчиков).

Анализируя полученные оценки, можно отметить следующее:

- если распознаваемый образ принадлежит к санкционированному пользователю, то в счетчике замечен существенный максимум в соответствующем разряде, а число других «ненулевых» разрядов невелико, если не брать во внимание разряды, содержащие по одной единице;

- номер класса, которому принадлежит распознаваемый образ, можно признать только тогда, когда число единичек в соответствующем разряде больше величины  $F/2$  для подобразов и/или больше  $3F/2$  для всех подобразов и при условии, что все максимумы принадлежат одному и тому же разряду счетчика.

На основе представленных в табл. 6.4.1 данных с учетом условий (6.4.5) можно теперь определить порог  $P_2$ . Очевидно, что он не может быть ниже величины 17. В этом случае все первые пять образов из рис. 6.4.6 будут идентифицированы как образы, принадлежащие 5-му классу. При таком пороге все остальные образы остаются неидентифицированными как санкционированные пользователи системы.

В заключение представим результат работы системы КД для двух случаев (показаны слева и справа на рис. 6.4.7).

В первом случае новый образ был однозначно идентифицирован на основе метода взаимной корреляции. Максимумы корреляции составили здесь 0,68, 0,61 и 0,38 соответственно для подобразов 3a, 3b, 3c. Два из них (подобразы 3a и 3b) превысили порог 0,6, что и определило окончание процесса идентификации.

Во втором случае максимумы корреляции были ниже порога 0,6, но все они соответствовали одному и тому же классу. Выполненная верификация на основе метрики  $L_2$  также подтвердила тот же самый класс, что представлено записью в форме {5 5 5 /// 5 5 5} под образом 2. При этом (в данном случае очевидно) суммарное состояние счетчика составило [0 0 0 0 27 0 0], что определило 100 %-ную верификацию.



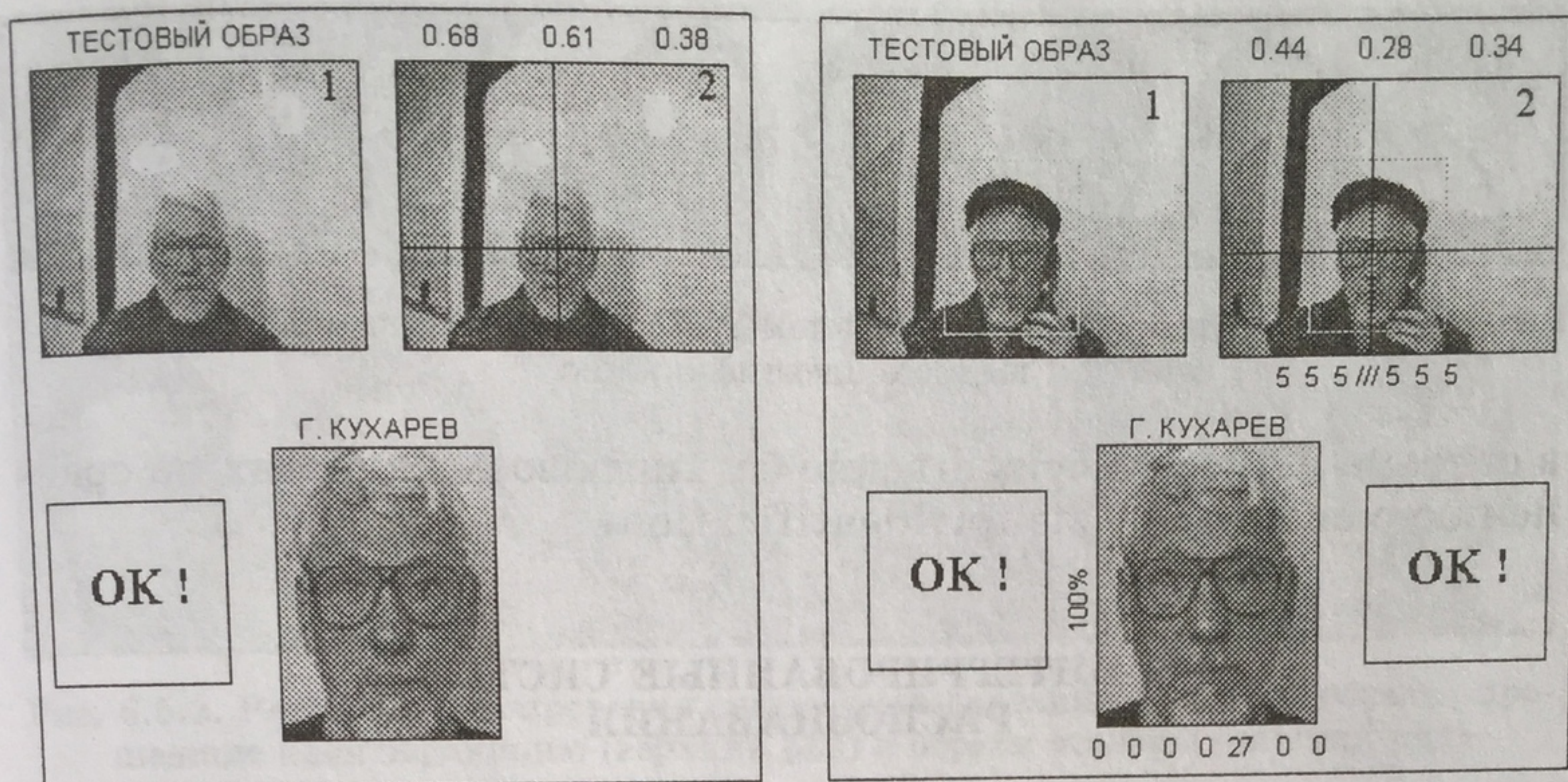


Рис. 6.4.7. Два варианта работы системы «Контроля доступа»

В завершение дадим качественную оценку эффективности работы системы «Контроль доступа». Система проверялась в двух режимах: «on-line» и «off-line». Во втором случае в качестве исходных данных использовался тестовый набор, представленный на рис. 6.3.5, а также вся база данных [Orl\*\*]. На рис. 6.4.8a показаны образы, прошедшие идентификацию (в режиме «on-line») как пользователи системы. На рис. 6.4.8b — ряд образов (из тестовых образов рис. 6.3.5), не прошедших идентификацию, поскольку при их предъявлении система не относила их к образам зарегистрированных пользователей.

Образы базы [Orl\*\*] также не прошли идентификацию. Эти результаты свидетельствуют о высоком качестве идентификации



Рис. 6.4.8a. Результат тестирования системы «Контроля доступа»: образы, прошедшие идентификацию и соответствующие им эталоны, хранящиеся в ВБ





Рис. 6.4.8b. Результат тестирования системы «Контроля доступа»: образы, не прошедшие идентификацию

в системе КД и о ее существенно больших возможностях по сравнению с системой «Visitor Identification».

### 6.5. ИНТЕГРИРОВАННЫЕ СИСТЕМЫ РАСПОЗНАВАНИЯ

Во всех рассмотренных выше системах распознавания используются биометрические характеристики лица, полученные из его двумерного представления. Эти системы просты в эксплуатации, обладают высоким быстродействием и, как было показано выше, могут обеспечивать требуемые показатели качества распознавания (например, FAR и FRR). Однако эти системы не защищены от несанкционированного доступа, поскольку могут быть обмануты плоскими изображениями лиц оригиналов (или фотомуляжей).

Покажем это экспериментально: попробуем распознать фотомуляжи, которыми будут непосредственно служить рисунки 6.3.6 и 6.4.3, «подставленные под камеру». Некоторые из этих фотомуляжей представлены на рис. 6.5.1.

Отметим две особенности этих образов — низкое качество (невысокое разрешение, неконтрастность изображений, наличие значительной шумовой составляющей, различие в освещении и т. д.)



Рис. 6.5.1. Фотомуляжи, полученные из рис. 6.3.6 и 6.4.3





Рис. 6.5.2. Результат тестирования систем распознавания «on-line»: образы, прошедшие идентификацию (верхний ряд) и образы эталоны (нижний ряд)

и зеркальное представление изображений лиц по отношению к лицам в базе данных (поскольку образом перед камерой является не лицо человека, а портрет лица).

Следует сразу отметить, что не все из представленных выше образов удалось распознать (рис. 6.5.2). Лучшие результаты при этом показала система «Контроль доступа». При этом те образы, которые были распознаны, приходилось предъявлять системе по несколько раз подряд (выполняя несколько попыток). Возможно, причиной этого было низкое качество фотомуляжей, особенности способа их получения и нестационарные условия их предъявления камере (лист с напечатанными образами держали в руках перед камерой), однако это также свидетельствует о некоторой защите тестируемых систем от обмана с помощью простых средств.

Ограничение числа попыток распознавания предъявляемых системе образов — один из способов защиты от обмана. Однако представленный на рис. 6.5.2 результат, свидетельствует о необходимости дополнительной защиты систем распознавания лиц от их обмана посредством примитивных фотомуляжей (зарегистрированных пользователей).

Один из вариантов защиты подобных систем состоит в их интегрировании с системами идентификации (аутентификации), использующими другие биометрические характеристики человека. Ясно, чем больше используется дополнительных биометрических образов, тем труднее изготовить необходимые муляжи. Варианты построения подобных систем рассмотрены, например, в [Hon98, Fri00].

В интегрированной системе [Hon98] дополнительной характеристикой является отпечаток пальца, а в системе [Fri00] — голос и движения губ. Метод распознавания отпечатков пальцев, используемый в [Hon98], аналогичен методу, описанному в разд. 2.1. Особенности подхода [Fri00], использующего три биометрические характеристики, были проанализированы в разд. 2.3.



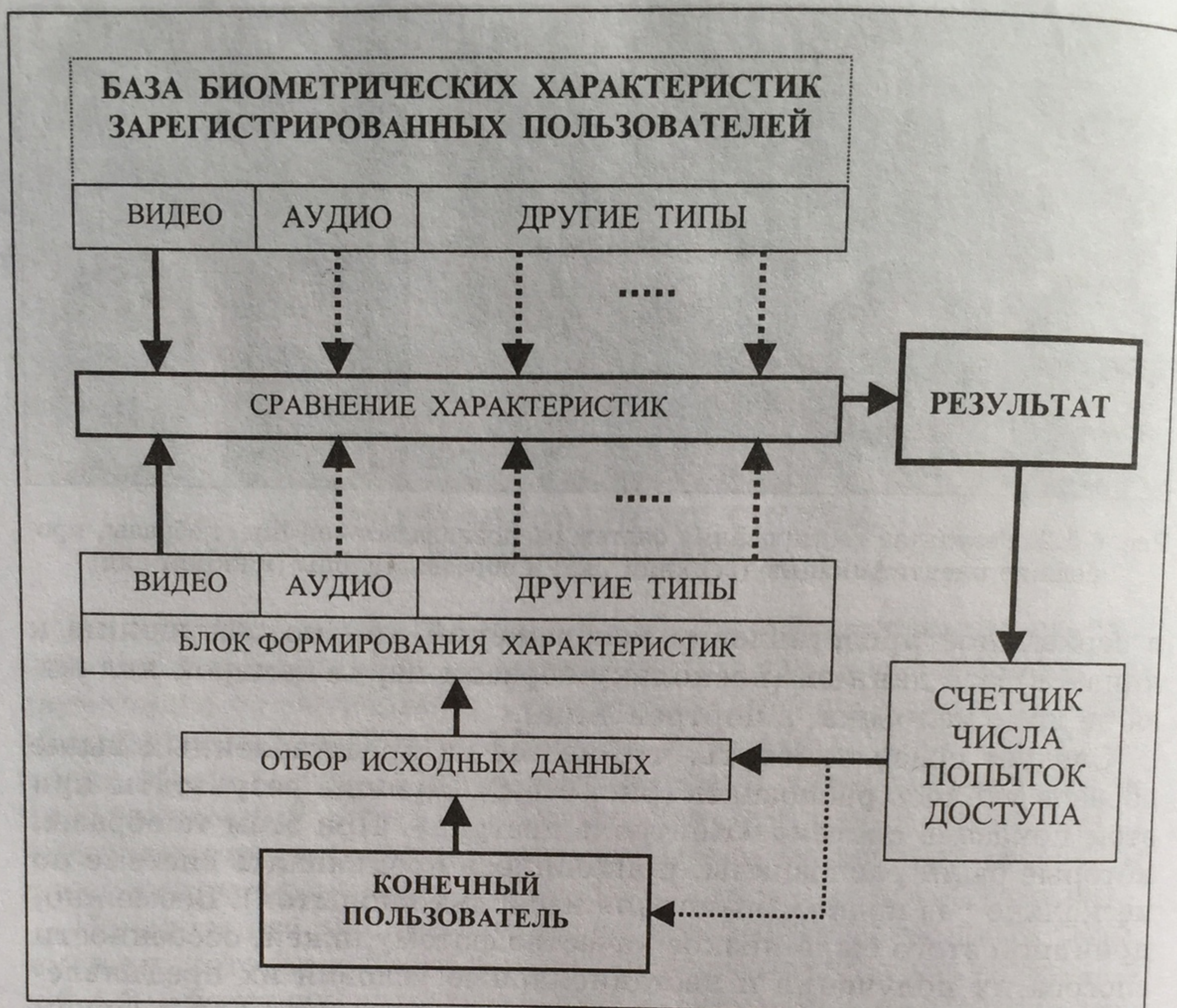


Рис. 6.5.3. Интегрированная биометрическая система

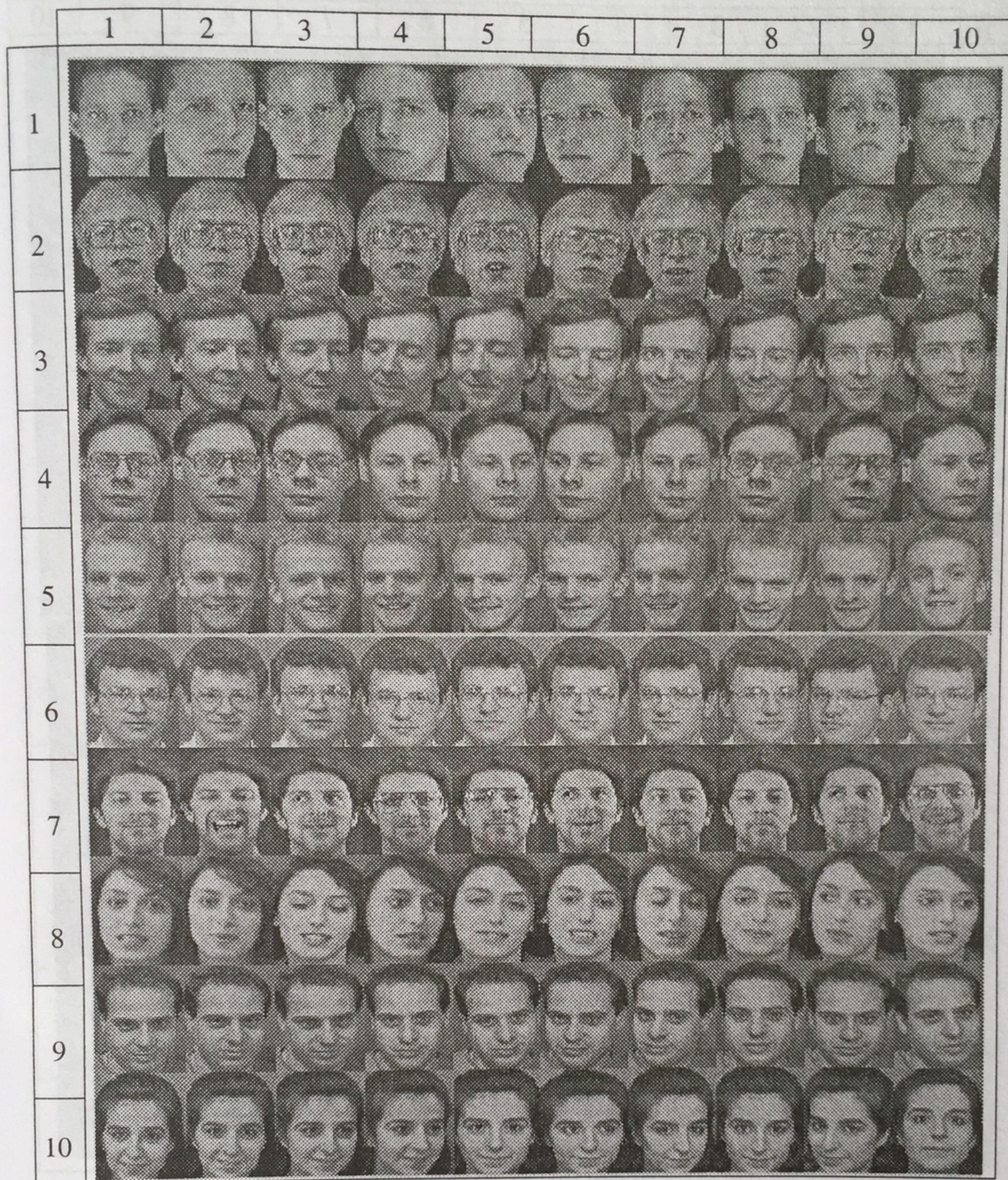
Общая структура интегрированных биометрических систем идентификации (аутентификации) человека представлена на рис. 6.5.3.

Обман таких систем становится практически невозможным из-за трудностей изготовления необходимых муляжей (имитаторов звуков или движения губ, например). Привлекательны и такие их достоинства, как простота в эксплуатации, относительно низкая стоимость и высокое быстроедействие. Это стало возможным благодаря использованию в них в качестве основного ядра системы распознавания человека по лицу.

Очевидно, что будущее систем идентификации личности человека связано с развитием именно интегрированных биометрических систем.

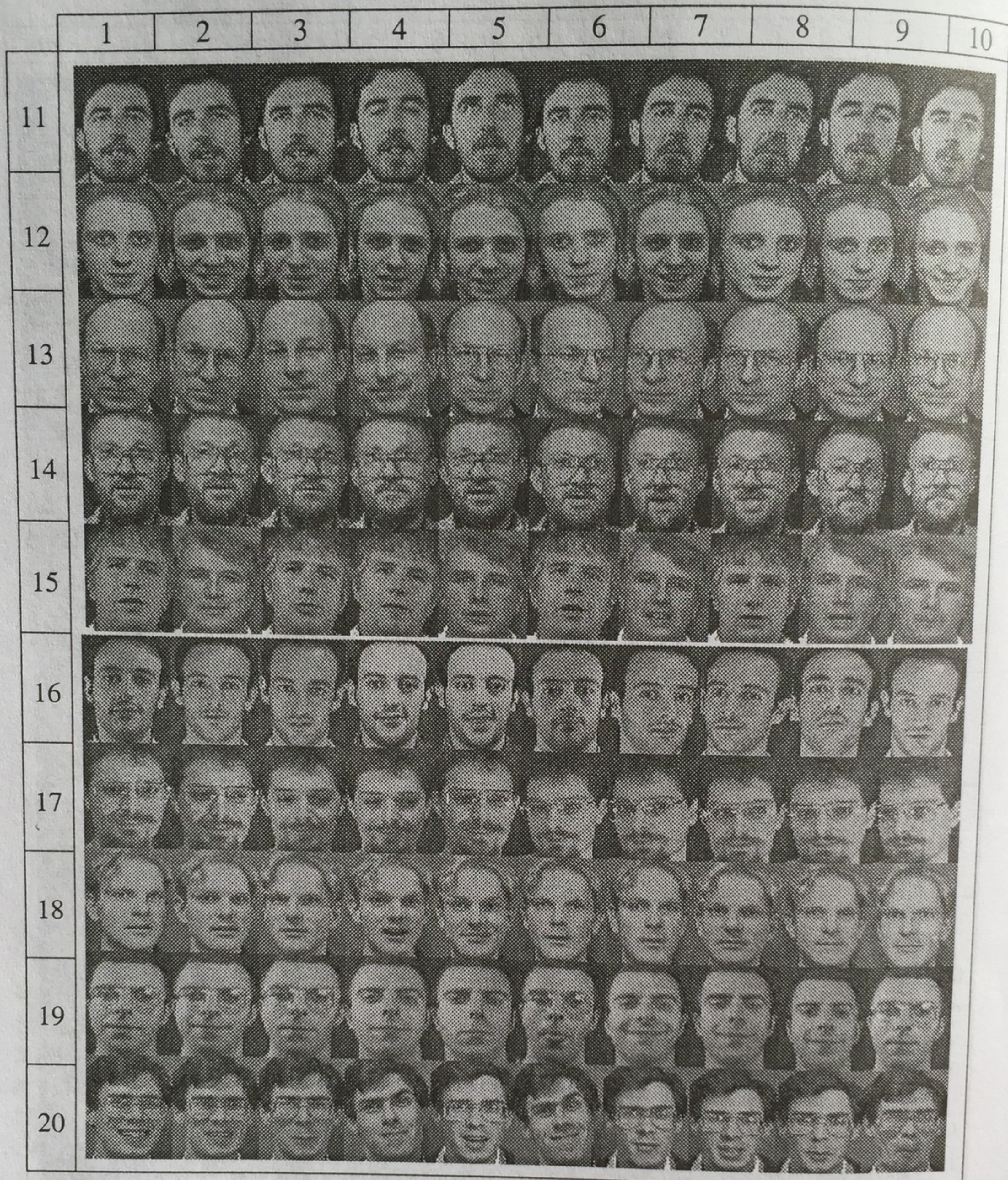


## Тестовая база данных (ORL Database of Faces)











































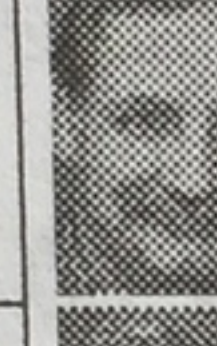





























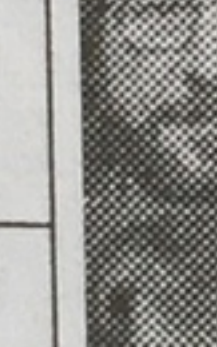









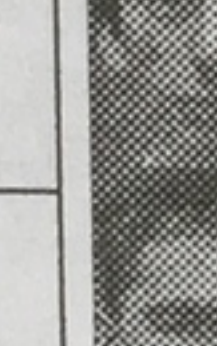









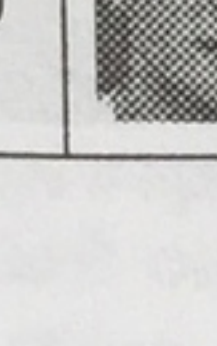
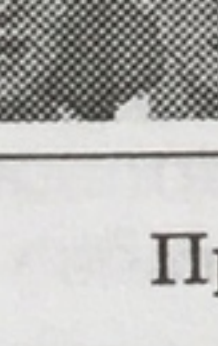
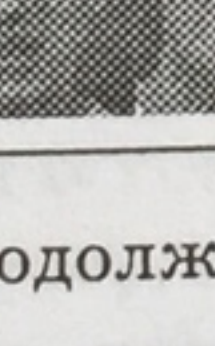
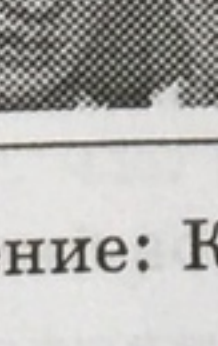
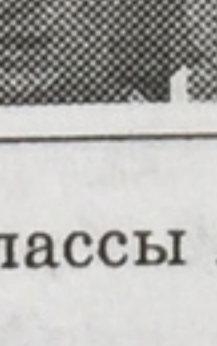
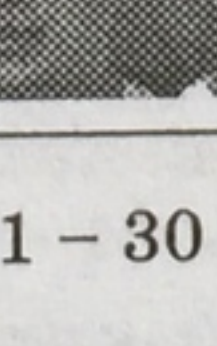
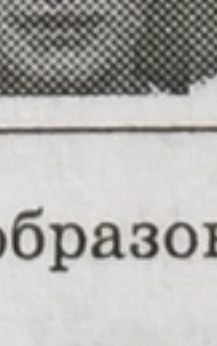
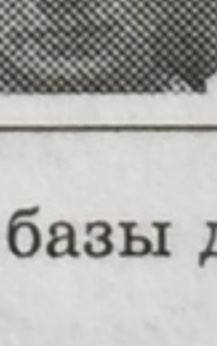
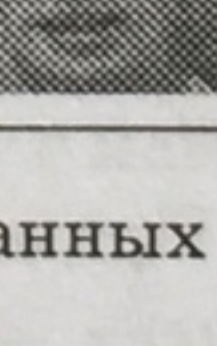
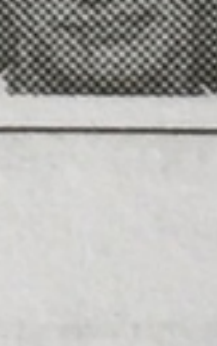
Первые 10 классов образов базы данных





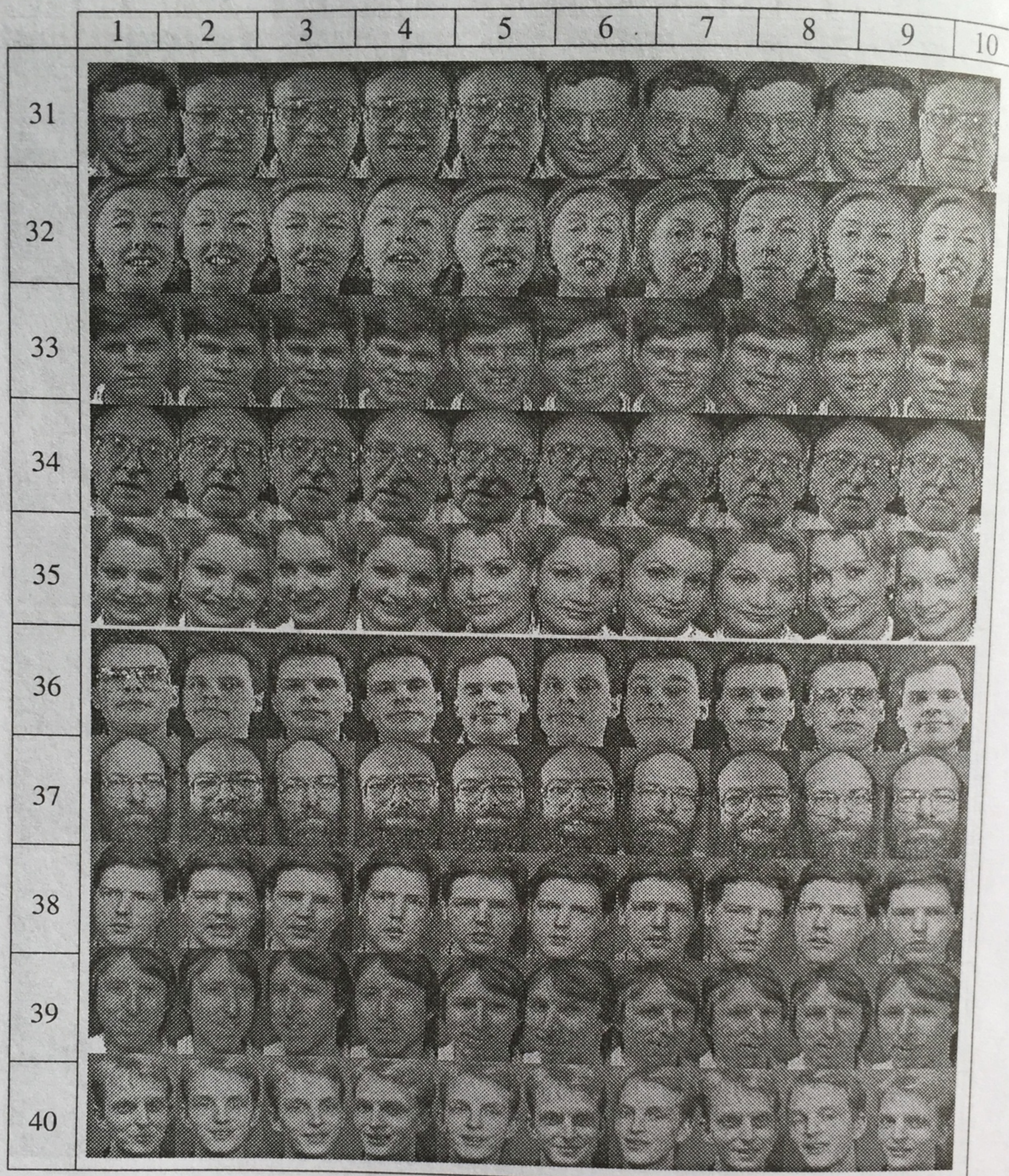
Продолжение: Классы 11 – 20 образов базы данных



|    | 1   | 2   | 3   | 4   | 5   | 6  | 7   | 8   | 9   | 10  |
|----|---|---|---|---|---|--|---|---|---|---|
| 21 |    |    |    |    |    |    |    |    |    |    |
| 22 |    |    |    |    |    |    |    |    |    |    |
| 23 |    |    |    |    |    |    |    |    |    |    |
| 24 |   |   |   |   |   |   |   |   |   |   |
| 25 |  |  |  |  |  |  |  |  |  |  |
| 26 |  |  |  |  |  |  |  |  |  |  |
| 27 |  |  |  |  |  |  |  |  |  |  |
| 28 |  |  |  |  |  |  |  |  |  |  |
| 29 |  |  |  |  |  |  |  |  |  |  |
| 30 |  |  |  |  |  |  |  |  |  |  |

Продолжение: Классы 21 – 30 образов базы данных





Продолжение: Классы 31 – 40 образов базы данных

A  
 Dete  
 ход  
 лиц  
 Fea  
 метр  
 точн  
 угол  
 A  
 (AFI  
 вает  
 писа  
 A  
 щие  
 биом  
 A  
 мето  
 боре  
 ких  
 вект  
 ствен  
 позв  
 бор  
 разм  
 ков  
 A  
 раже  
 ции  
 зави  
 A  
 опре  
 как  
 лени  
 A  
 пред  
 систе  
 ки ут  
 торы  
 B  
 ные  
 пред  
 струк  
 ходн  
 ной  
 B  
 щая  
 мину



**Автоматический поиск области лица на изображениях (Face Detection)** — автоматическое определение координат подобраза на исходном изображении, включающем собственно лицо, и выделение образа лица из исходного изображения.

**Автоматический поиск элементов лица человека (Automatic Facial Features Detection)** — автоматическое определение координат антропометрических точек лица таких, например, как уголки глаз, центры глаз, точка пересечения линии симметрии лица и линии глаз, кончик носа, уголки губ, центр губ и т.д.

**Автоматизированная система идентификации отпечатка пальца (AFIS)** — специализированная биометрическая система, которая сравнивает заданное изображение пальца с эталонами изображений пальца, записанными в базе данных.

**Актуальный эталон (Reference Template)** — данные, представляющие биометрические характеристики зарегистрированного пользователя биометрической системы.

**Анализ главных компонент (Principal component analysis — PCA)** — метод, позволяющий *выявить* основные изменчивости в обучающем наборе исходных данных и *описать* эти изменчивости при помощи нескольких наибольших собственных чисел и соответствующих им собственных векторов; ортогональное преобразование, выполненное в базисе этих собственных векторов (иначе преобразование Карунена—Лоэва или ПКЛ), позволяет трансформировать исходные признаки (представляющие набор исходных данных) в пространство признаков существенно меньшей размерности; этот процесс называется редукцией пространства признаков и именно в результате ПКЛ определяются главные компоненты.

**Анализ изображений (Image Analysis)** — процесс выделения из изображения неизобразительной информации, а также описания/интерпретации и представления изображения с помощью его компонентов, свойств и зависимостей.

**Анализ эмоций по выражению лица (Facial Expressions Analysis)** — определение типов выражения лица, которые можно интерпретировать как человеческие эмоции (например, страх, гнев, отвращение, грусть, удивление, счастье).

**Аутентификация (Authentication)** — подтверждение идентичности представленного образа одноименному эталону, хранящемуся в памяти системы распознавания; аутентификация реализуется на основе проверки утверждения о том, что представленный образ есть именно тот, за который он себя «выдает».

**База данных, исходная база данных (Database)** — структурированные или каталогизированные данные, представляющие тестовые образы, предназначенные для проверки эффективности систем распознавания; структурированные или каталогизированные данные, определяющие исходные образы санкционированных пользователей для каждой конкретной системы распознавания.

**База данных минуций (Minutiae Database)** — база данных, содержащая эталонные отпечатки пальцев в форме вектора минуций. Каждая минуция определяется координатами  $X$  и  $Y$ , углом ориентации папилляр-



ной линии в данной координате и дополнительно типом минуции (например, окончание или раздвоение линии). Каждый отпечаток пальца может быть представлен 50 – 70 минуциями.

**Банк паролей (Password Bank)** — база данных, содержащая имя пользователя, его пароль и другую персональную информацию о нем; все эти данные предназначены для того, чтобы допустить пользователя до системы и реализовать его идентификацию и/или верификацию.

**Биометрика (Biometrics)** — область знаний, представляющая методы измерения физических характеристик (или персональных поведенческих черт) человека и методы их сравнения с аналогичными данными в базе данных для целей идентификации или аутентификации человека.

**Биометрическая система (Biometric Systems)** — автоматизированная система, способная реализовать функции: фиксации биометрической выборки от конечного пользователя; извлечения биометрических данных из той выборки; сравнения биометрических данных с одним или большим количеством эталонов; принятия решения о том, как хорошо они соответствуют; и индикации о том, была или нет достигнута идентификация или проверка идентичности.

**Биометрическая характеристика (Biometric Characteristic)** — измеримая *физическая характеристика* человека, используемая в процессе его проверки на подлинность или идентичность санкционированному пользователю.

**Поведенческая биометрическая характеристика (Behavioral Biometric Characteristic)** — персональная *поведенческая черта* человека, используемая в процессе его проверки на подлинность или идентичность санкционированному пользователю.

**Биометрические технологии (Biometric Technology)** — отрасль (науки и производства), представляющая методы и технические средства получения и использования биометрических данных человека в целях его идентификации (верификации, аутентификации или распознавания). При этом в биометрических технологиях используются как *физические биометрические характеристики* человека — отпечатки пальцев, геометрия руки, изображения радужной оболочки и сетчатки глаза, голос, видео- и термоизображения лица, подпись (а в последнее время изображения ушей, отпечатки губ и запах человека), так и *поведенческие биометрические характеристики*, например, манера работы на клавиатуре компьютера, динамика написания рукописного текста, стиль и манера походки.

**Верификация (Verification)** — дополнительная проверка результата распознавания с целью установления истинности этого результата; верификация реализуется аналогично аутентификации.

**Виртуальная база данных (Virtual Database)** — база данных, содержащая образы-эталон в той форме, в которой они непосредственно используются системой (программой, алгоритмом) распознавания.

**Время захвата (Live Capture)** — период интеракции между конечным пользователем и биометрической системой с целью получения биометрической выборки.

**Время ответа/ Время обработки (Response Time/Processing Time)** — период, необходимый, чтобы биометрическая система выработала решение по идентификации или проверке биометрической выборки.

**Время регистрации (Enrollment Time)** — время, которое должна затратить биометрическая система на получение биометрического образа (эталона)



пользователя, включая фиксацию биометрической выборки от пользователя системы и извлечение биометрических данных (образа) из той выборки.

**Глобальные признаки отпечатка пальца (Global Fingerprint Features)** — центры специфических участков отпечатка (например, дельты, спирали, петли и т. д.).

**Зарегистрированный пользователь (Enrollee)** — человек, имеющий актуальный биометрический эталон в базе данных (или на файле) системы.

**Захват данных (Capture)** — метод получения биометрической информации (данных, выборки) от конечного пользователя.

**Идентификация биометрическая (Biometric Identification)** — процедура, базирующаяся на технологии распознавания образов, и предназначенная для однозначного определения личности человека на основе его биометрических характеристик при сравнении их с заданными эталонами.

**Идентификация образов (Pattern Identification)** — процесс вынесения решения об образе или собственно решение, которые приписывают этому образу однозначное название.

**Извлечение (Extraction)** — процесс преобразования биометрической выборки в биометрические данные и приведение их к некоторой стандартизированной форме так, чтобы можно было их сравнивать с заданным эталоном.

**Изображение цифровое (Digital Image)** — двумерное изображение, представленное в цифровой форме в виде числовой матрицы, каждый элемент которой может принимать  $2^L$  значений яркости (обычно,  $1 \leq L \leq 8$ ).

**Классификация образов (Pattern Classification)** — процесс вынесения решения об образе или собственно решение, которое указывает лишь на принадлежность образа к определенному классу образов.

**Кластеризация (Clusterization)** — процесс (или результат) группирования исходных образов в совокупности — кластеры — на основе одинаковых или близких значений некоторых признаков, присущих исходным образам.

**Конечный пользователь (End User)** — человек, который взаимодействует с биометрической системой, чтобы зарегистрироваться или проверить идентичность.

**Конкатенация (Concatenation)** — процесс (или результат) объединения в одну последовательность отдельных подпоследовательностей знаков или символов, например, конкатенация строк или столбцов числовой матрицы в соответствующий вектор. В задачах распознавания двумерных образов, задаваемых соответствующими матрицами, конкатенация матрицы по строкам или столбцам переводит этот образ из двумерного представления в одномерное (векторное).

**Линейный дискриминантный анализ (Linear Discriminant Analysis — LDA)** — процедура анализа и обработки исходных данных в задачах распознавания образов (в качестве исходных данных выступают признаки, характеризующие образы в отдельных классах) с целью улучшения их кластеризации. LDA основан на линейном преобразовании в базисе собственных функций, который получается в результате решения задачи на собственные значения для матрицы ковариации. Последняя, в свою очередь, представляет дисперсионную изменчивость исходных данных как внутри классов образов, так и между классами образов.

**Ложный отказ (False Rejection)** — случай, когда биометрическая система не подтверждает истинность зарегистрированного пользователя.



**Минуции (Minutiae)** — точки, соответствующие разрывам, окончанием и раздвоениям бороздок папиллярных линий на образах отпечатка пальца; минуции определяют локальные признаки отпечатка пальца и описаны в эталоне отпечатка.

**Модуль идентификации отпечатка пальца (Fingerprint Identification Unit — FIU)** — биометрическая система, способная к получению, сохранению и сравнению данных отпечатков пальцев для проверки идентичности индивидуума.

**Непосредственное чтение отпечатка пальца (DFR)** — устройство, предназначенное для непосредственного сканирования изображения пальца из пальца индивидуума.

**Один из многих (One-to-Many)** — поиск по заданному образу подобных образов в базе данных с тем, чтобы определить существует или нет кандидат в базе данных (синоним идентификации).

**Один из одного (One-to-One)** — установление идентичности заданного образа некоторому образу-эталону (синоним верификации).

**Персональный номер идентификации (Personal Identification Number — PIN)** — номер с четырьмя цифрами, вводимый пользователем в систему по ее запросу, и используемый как код доступа к специфической области системы; использование PIN определяет один из методов защиты доступа к системе.

**Пиксель (Picture cell, pixel)** — наименьший элемент цифрового изображения, который характеризуется собственным значением (уровнем) яркости.

**Попытка (Attempt)** — подача биометрической выборки к биометрической системе для идентификации или верификации. Биометрическая система может потребовать больше одной попытки идентификации или проверки.

**Порог (Threshold)** — предельное значение некоторого заданного параметра, определяющее условие принятия решения, например, в процедуре верификации.

**Преобразование Карунена—Лоэва (ПКЛ)** — ортогональное преобразование исходных данных, выполняемое в базисе собственных функций, который получается в результате решения задачи на собственные значения для матрицы ковариации, определяемой для исходных данных. Матрица ковариации, в свою очередь, представляет дисперсионную изменчивость исходных данных.

**Признак (Feature)** — компоненты образа (например, выделенный объект, текстура), атрибуты образа (например, его размер, форма) или характеристики образа (например, яркость, спектр), используемые в распознавании образов; в общем случае признак определяется той категорией, посредством которой он выражен; известны и используются три категории признаков — физические, структурные и математические.

**Программа технологии распознавания лиц (Face Recognition Technology/FERET)** — комплексная разработка (агентства DARPA и исследовательской лаборатории армии США), предназначенная для оценки эффективности предлагаемых алгоритмов распознавания лиц и стимулирования развития наилучших из них.

**Распознавание лиц (Face Recognition)** — раздел биометрики, представляющий методы и средства решения задач распознавания и наблюдения человека (идентификацию или классификацию) по изображению



ям его лица, а также интерпретации действий человека — его намерений и поведения — через идентификацию его эмоционального состояния и окружающих его людей.

**Распознавание образов (Pattern Recognition)** — процесс, складывающийся из двух частей: 1 — определения и извлечения признаков; 2 — принятия решения о принадлежности предъявленного образа (представленного в виде совокупности признаков) к одному из известных классов (при этом решается либо задача идентификации, либо задача классификации).

**Регистрация пользователя (Enrollment)** — процесс сбора биометрических данных от пользователя, включающий в себя: фиксацию биометрической выборки от пользователя системы, извлечение биометрических данных (образа) из той выборки, сохранение текущих (актуальных) биометрических данных; в результате регистрации получают исходные биометрические данные (исходный образ/образы), представляющие конкретного индивидуума.

**Собственные лица (Eigen Faces)** — элементарный образ лица, реконструированный на основе главных компонент и обратного преобразования Карунена—Лоэва.

**Сопоставление с эталоном (Template Matching)** — процесс сравнения биометрической выборки с эталоном и оценка уровня подобия между ними.

**Сравнение (Comparison)** — процесс сравнения биометрических данных конкретного человека с заданным эталоном на основе некоторой метрики.

**Станция регистрации (Enrollment Station)** — автоматизированное рабочее место, в котором биометрика индивидуума (отпечаток пальца, голос, изображение лица и т. д.) и его персональная информация (имя, адрес, и т. д.) могут быть введены в систему биометрической идентификации.

**Уровень ошибочных отказов (False Reject Rate — FRR)** — вероятность того, что биометрическая система не идентифицирует зарегистрированного пользователя или не подтверждает его подлинности.

**Уровень ошибочных подтверждений (False Acceptance Rate — FAR)** — вероятность того, что биометрическая система ошибочно идентифицирует несанкционированного (незарегистрированного) пользователя и/или подтверждает его подлинность.

**Цифровая подпись (Digital Signature)** — образ подписи, представленный в цифровой зашифрованной форме, не позволяющей чтение/интерпретацию подписи без использования соответствующего дешифрирующего ключа.

**Шифрование (Encryption)** — преобразование открытого текста в практически не читаемую форму с использованием математического преобразования.

**Эталон отпечатка пальца (Fingerprint Template)** — описание всех обнаруженных минуций (деталей) в образе отпечатка пальца; эталон содержит координаты  $X$  и  $Y$  каждой из минуций, ее тип и наклон, которые характеризуют конкретный отпечаток пальца и используются для сравнения с отпечатками пальцев всех прочих кандидатов.

\* — Подготовлен на основе Biometric Glossary [Bio\*\*].



## ЛИТЕРАТУРА

- [Бар90]. Барабашева Ю. М., Девяткова Г. Н., Микешина Н. Г. и др. Компьютерная биометрика. — М.: Из-во МГУ, 1990. — 232 с.
- [Бов97]. Бовбель Е. И., Кухарев Г. А., Кухарчик П. Д. Принципы построения и опыт распознавания речи для мультимедиа технологии. — Мат-лы первой Международной конфер. «Applications of Computer Systems», Szczecin, 1994, 107 – 116 с.
- [Вел88]. Вельховер Е. С. и др. Иридодиагностика. — М.: Медицина. — 240 с.
- [Вел98]. Вельховер Е. С., Вершинин Б. В. Тайные знаки лица. — М.: Изд-во Российского университета дружбы народов, 1998. — 328 с.
- [Гла97]. Главные компоненты временных рядов: Метод «Гусеница»/Под ред. Данилова Д. и Жиглявского А. — СПб.: СПГУ, 1997. — 308 с.
- [Гол99]. Голуб Дж., Ван Лоун Ч. Матричные вычисления: Пер. с англ. — М.: Мир, 1999. — 548 с.
- [Гон97]. Гоникман Э. И. Искусство диагностики по лицу. — Минск.: Сантана, 1997. — 473 с.
- [Даг83]. Дагман Э. Е., Кухарев Г. А. Быстрые дискретные ортогональные преобразования. — Новосибирск, Наука, 1983. — 233 с.
- [Евс97]. Евстигнеев В., Черномордик О. Биометрическая идентификация личности//Банковские технологии. — 1997. — №1.
- [Зав98]. Завгородний В. В. Биометрия: новые технологии идентификации личности//Банковские технологии. 1998. — № 10.
- [Ках98]. Каханер Д., Моулер К., Нэш С. Численные методы и математическое обеспечение: Пер. с англ. — М.: Мир, 1998. — 575 с.
- [Кух97]. Кухарев Г., Borawski M. Методика обнаружения объектов на 2D-изображениях посредством Фурье-описания контуров объектов и согласованной фильтрации. — Proceedings of Fourth International Conference Pattern Recognition and Information Processing (PRIP'97), 20 – 22 May 1997, Minsk, Republic of Belarus, vol. 2, p. 98 – 104.
- [Кух99a]. Кухарев Г. Приближенный метод оценки угла поворота изображений по их спектрам. — Мат-лы II Международной науч.-техн. конфер. «Современные методы цифровой обработки сигналов в системах измерения, контроля, диагностики и управления». 24 – 27 июня 1998 (Опубликовано в 1999 г.), Минск, с. 39 – 43.
- [Кух99b]. Кухарев Г. Распознавание лица человека на основе одного эталона. — Proceedings of Fifth International Conference Pattern Recognition and Information Processing (PRIP'99), 18 – 20 May 1999, Minsk, Republic of Belarus, vol. 2, p. 15 – 20.
- [Кух99c]. Кухарев Г., Soldek J. Распознавание образов с использованием инвариантов геометрических преобразований. — Мат-лы II Международной науч.-техн. конфер. «Современные методы цифровой обработки сигналов в системах измерения, контроля, диагностики и управления». 24 – 27 июня 1998 (Опубликовано в 1999 г.), Минск, с. 44 – 48.
- [Кух99d]. Кухарев Г. Реализация и анализ простой системы распознавания лиц человека. — Proceedings of Fifth International Conference Pattern Recognition and Information Processing (PRIP'99), 18 – 20 May 1999, Minsk, Republic of Belarus, vol. 2, p. 63 – 67.
- [Лев99]. Левшинов А. 12 сторон ладони. — СПб: Питер, 1999. — 224 с.
- [Лео99]. Леонов В. П. История биометрики и ее применение в России//Международный журнал медицинской практики, 1999. — Вып. 4. — С. 7 – 19.
- [Лев22]. Левитский Г. Элементы биометрики. — М., 1922, — 78 с.
- [Мат00]. Матвиенко Р. Когда палец становится ключом//Магия ПК. — 2000. — № 10. — С. 30 – 32.



[Отс00]. Открытые системы. — 2000. — № 3. — С. 17 – 33.

[Пав86]. Павлидис Т. Алгоритмы машинной графики и обработки изображений: Пер. с англ. — М.: Радио и связь, 1986. — 400 с.

[Пан00]. Панканти Ш., Боле Р., Джейн Э. Биометрия: будущее идентификации//Открытые системы. — 2000. — № 3. — С. 17 – 20.

[Пен00]. Пентланд А. Распознавание лиц для интеллектуальных сред//Открытые системы. — 2000. — № 3. — С. 17 – 20.

[Пер98]. «Перст судьбы»//КП. — 1998. — 6 ноября. — С. 19

[Пот99]. Потемкин В. Г. Система инженерных и научных расчетов MATLAB 5.x. — В 2-х т. — М.: Диалог. МИФИ, 1999. — 366 + 304 с.

[Поо98]. По отпечаткам пальцев//Разбор. — 1998. — №11.

[Прэ82]. Прэтт У. Цифровая обработка изображений: Пер. с англ. — М.: Мир. 1982. — Кн. 1 и 2.

[Руд00]. Рудаков П. И., Сафонов В. И. Обработка сигналов и изображений. MATLAB 5.x. — М.: Диалог. МИФИ, 2000. — 416 с.

[Сам98]. Самаль Д., Старовойтов В. Подходы и методы распознавания людей по фотопортретам. — Минск. Препринт № 8, ИТК НАНБ, 1998. — 53 с.

[Тел98]. Телемедицина: Новые информационные технологии на пороге XXI века. — СПб.: РАН. СПб. институт информатики и автоматизации, 1998. — 489 с.

[Теп95]. Теппервайн К. Лицо — зеркало здоровья. Патофизиогномика для всех/Пер. с нем. — Изд-во «ПитерПресс», 1995. — 184 с.

[Фил00]. Филлипс П. Д., Мартин Эл., Уилсон С. Л., Пржибоски М. Введение в оценку биометрических систем//Открытые системы. — 2000. — № 3. — С. 21 – 27.

[Янш94]. Яншин В. В. Анализ и обработка изображений: Принципы и алгоритмы. — М.: Машиностроение. 1994. — 112 с.

[Adi97]. Adini Y., Moses Y., Ullman S. Face Recognition: The Problem of Compensating for Changes in Illumination Direction//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — N 7 (July). — P. 721 – 731.

[Alf99]. Alferez R., Wang Y.-F. Geometric and Illumination Invariants for Object Recognition//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1999. — Vol. 21. — N 6 (June). — P. 505 – 535.

[Ame\*\*]. American National Standards Institute (ANSI), Fingerprint Identification. — Data Format for Information Interchange.

[Bel97]. Belhumeur P. N., Hespanha J. P., Krejgman D. J. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — N 7 (July). — P. 711 – 720.

[Ben98a]. Ben-Arie J., Nandy D. A Volumetric/Iconic Frequency Domain Representation for Objects With Application for Pose Invariant Face Recognition//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1998. — Vol. 20. — N 5 (May). — P. 449 – 457.

[Ben98b]. Ben-Arie J., Wang Z. Pictorial Recognition of Objects Employing Affine Invariance in the Frequency Domain//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1998. — Vol. 20. — N 6 (June). — P. 604 – 618.

[Bio\*\*]. <http://www.biometrics.org/>, <http://www.iosoftware.com/>

[Bor97]. Borawski M., Chimiak J., Kuchariew G. The Method of Personal Identification Based on the Iris Image Recognition. Proceedings of the Fourth International Conference Applications of Computer Systems, November 1997, Szczecin, Poland, p. 78 – 84.

[Bru93]. Bruneli R., Poggio T. Face Recognition: Features versus Templates//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1993. — Vol. 15. — N 10 (October). — P. 1042 – 1052.

[Bry99]. Brylka T., Ponikowski T. Image Preprocessing for Face Identification System. Proceedings of the Sixth International Conference Advanced Computer Systems. ACS'99, November 1999, Szczecin, Poland, p. 79 – 86.



[Cap99]. Cappelli R., Lumini A., Maio D. Fingerprint Classification by Directional Image Partitioning//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1999. — Vol. 21. — N 5 (May). — P. 402 – 421.

[Com00]. Computer, February 2000, p. 46 – 80.

[Che94]. Chen D., Deconinck F. Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1994. — Vol. 16. — N 12 (December). — p. 1156 – 1168.

[Don99]. Donato G., Bartlett M. and others. Classifying Facial Actions//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1999. — Vol. 21. — N 10 (October). — P. 974 – 989.

[Ess97]. Essa I., Pentland A. Coding, Analysis, Interpretation and Recognition of Facial Expressions//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — N 7 (July). — P. 757 – 763.

[Gol97]. Golfarelli M., Maio D., Maltoni D. On the Error-Reject Trade-Off in Biometrics Verification Systems//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — N 7 (July). — P. 786 – 796.

[Gou96]. Goudail F., Lange E., Iwamoto T., Kyuma K., Otsu N. Face Recognition System Using Local Autocorrelations and Multiscale Integration//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1996. — Vol. 18. — N 10 (October). — P. 1024 – 1028.

[Fin00]. Fingerprint Verification Competition 2000, www.csr.unibo.it/boilab.

[Fri00]. Frischholz R. W., Dieckmann U. BioID: A Multimodal Biometric Identification System//IEEE Computer. — 2000 (February). — P. 64 – 68.

[Hon98]. Hong Lin, Jain Anil. Integrating Faces and Fingerprints for Personal Identification//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1998. — Vol. 20. — N 12 (December). — P. 1295 – 1307.

[Jai97a]. Jain A. K., Hong L., Pankanti S. and Bolle R. An Identity-Authentication System Using Fingerprints//Proceedings of IEEE. — 1997. — Vol. 85. — N 9 (September). — P. 1365 – 1388.

[Jai97b]. Jain A. K., Hong L., Bolle R. On-line Fingerprint Verification//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — N 4 (April). — P. 302 – 313.

[Jai00]. Jain A., Hong L., Pankanti S. Biometric Identification. Communications of the ACM. — 2000. — Vol. 43. — N 2 (February). — P. 91 – 98.

[Kar98]. Karpouzis K., Votsis G., Tsapatsoulis N., Kollias S. Compact 3D Model Generation Based on 2D views of Human Faces: Application to Face Recognition. Machine GRAPHICS & VISION. — 1998. — Vol. 7. — N 1/2. — P. 75 – 85.

[Kir90]. Kirby M., Sirovich L. Application of the Karhunen-Loeve Procedure for Characterization of Human Face//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1990. — Vol. 13. — N 1 (January). — P. 103 – 108.

[Kuc97a]. Kuchariew G. Direct and Inverse Transformations of Spectra from Cartesian into Polar Systems of Coordinates. — Proceedings of Fourth International Conference Pattern Recognition and Information Processing (PRIP'97), 20 – 22 May 1997, Minsk, Republic of Belarus. — Vol. 1. — P. 92 – 102.

[Kuc98a]. Kuchariew G. Przetwarzanie i analiza obrazów cyfrowych. — Skrypt Instytutu Informatyki PS, Wyd. «Informa», Szczecin, 1998, 138 s.

[Kuc98b]. Kuchariew G., Borawski M., Forczmański P. Badanie metod obróbki obrazów cyfrowych w zastosowaniu do rozpoznawania twarzy. — Materiały III Sesji Naukowej Informatyki. INFORMA, Szczecin, 1998, s. 173 – 180.

[Kuc98c]. Kuchariew G., Forczmański P. New Algorithms of Spectra Transformations from Cartesian into Polar System of Coordinates for Image Recognition and Registration. — Proceedings of The Fifth International Conference on Advanced Computer Systems ACS 98, November 19 – 20, 1998, Szczecin, Poland, p. 56 – 63.

[Kuc99a]. ...  
nia. — Materia

[Kuc99b]. ...  
Realtime Face  
Sixth Interna  
Szczecin, Polan

[Kuc00a]. ...  
Identification.  
Systems. ACS'2

[Kuc00b]. ...  
of Face Recogn  
Computer Syste

[Kuc00c]. ...  
Algorithms. —  
Systems. ACS'2

[Kuc00d]. ...  
obrazowych dla  
INFORMA, Szc

[Kuc01a]. ...  
Identification a  
Pattern Recogn  
Republic of Bel

[Kuc01b]. K...  
Dimensionality  
Sixth Internatio  
18 – 20 May 200

[Lam98]. La...  
Based on a Singl  
ce. — 1998. —

[Lan97]. Lan...  
of Face Images  
Intelligence. —

[Liu00]. Liu...  
Recognition//IE  
Vol. 22. — N 6

[Mul98]. Mu...  
Skarbka. AOW

[Neg00]. Ne...  
Use//Computer.

[Orl\*\*]. AT&...  
ftp.uk.research.

[Pen00]. Pen...  
IEEE Computer.

[Phi98]. Phi...  
Recognition Algor

[Rat96]. Rat...  
for Lagre Finge  
Intelligence. —

[Rau92]. Rau...  
Parametric Polar  
on Machine Vision

[Row98]. Row...  
IEEE Trans. Patte  
nuary). — P. 23 –



[Kuc99a]. Kuchariew G. Inwarianty Obrazów Konturowych dla Zadań Rozpoznawania. — Materiały IV Sesji Naukowej Informatyki. INFORMA, Szczecin 1999, s. 51 – 58.

[Kuc99b]. Kuchariew G., Forczmański P., Kraszewski P., Bryłka T., Ponikowski T. Realtime Face Recognition System for Personal Identification. — Proceedings of the Sixth International Conference Advanced Computer Systems. ACS'99, November 1999, Szczecin, Poland, p. 58 – 63.

[Kuc00a]. Kuchariew G., Tujaka A. Real-time Face Recognition System for Visitor Identification. Proceedings of the Seventh International Conference Advanced Computer Systems. ACS'2000, October 2000, Szczecin, Poland, p. 516 – 522.

[Kuc00b]. Kuchariew G., Tsishkou D., Labuda A. Multistream Approach to Problem of Face Recognition. — Proceedings of the Sixth International Conference Advanced Computer Systems. ACS'2000, October 2000, Szczecin, Poland, p. 558 – 563.

[Kuc00c]. Kuchariew G., Boubel E., Tsishkou D. Face Identification Using Genetic Algorithms. — Proceedings of the Sixth International Conference Advanced Computer Systems. ACS'2000, October 2000, Szczecin, Poland, p. 564 – 579.

[Kuc00d]. Kuchariew G., Forczmański P. Kompresja i uporządkowanie danych obrazowych dla zadań, rozpoznawania. — Materiały V Sesji Naukowej Informatyki. INFORMA, Szczecin 2000, s. 15 – 21.

[Kuc01a]. Kuchariew G., Tujaka A. Pattern Recognition Methods for Visitor Identification and Access Control. — Proceedings of Sixth International Conference Pattern Recognition and Information Processing (PRIP'01), 18 – 20 May 2001, Minsk, Republic of Belarus. — Vol. 1. — P. 19 – 34.

[Kuc01b]. Kuchariew G., Forczmański P. Hierarchical Method of Reduction of Features Dimensionality for Image Recognition and Graphical Data Retrieval. — Proceedings of Sixth International Conference Pattern Recognition and Information Processing (PRIP'01), 18 – 20 May 2001, Minsk, Republic of Belarus. — Vol. 1. — P. 51 – 71.

[Lam98]. Lam K.-M., Yan H. An Analitic-to-Holistic Approach for Face Recognition Based on a Single Frontal View//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1998. — Vol. 20. — N 7 (July). — P. 673 – 686.

[Lan97]. Lanitis A., Taylor Ch. J., Cootes T. F. Automatic Interpretation and Coding of Face Images Using Flexible Models//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 17. — N 7 (July). — P. 743 – 755.

[Liu00]. Liu C., Wechsler H. Evolutionary Pursuit and Its Application to Face Recognition//IEEE Trans. Pattern Analysis and Machine Intelligence. — 2000. — Vol. 22. — N 6 (June). — P. 570 – 582.

[Mul98]. Multimedia: Algorytmy i Standardy Kompresji./pod redakcją Władysława Skarbka. AOW PLJ, Warszawa, 1998.

[Neg00]. Negin M. and others. An Iris Biometric System for Public and Personal Use//Computer. — 2000 (February). — P. 70 – 75.

[Orl\*\*]. AT&T Laboratories Cambridge — The ORL Database of Faces, [ftp://ftp.uk.research.att.com/pub/data/att\\_faces.zip](ftp://ftp.uk.research.att.com/pub/data/att_faces.zip)

[Pen00]. Pentland A., Choudbury T. Face Recognition for Smart Environments//IEEE Computer. — 2000 (February). — P. 50 – 55.

[Phi98]. Phillips et al. The Feret Database and Evaluation Procedure for Face Recognition Algorithms//Image and Vision Computing. — 1998. — May. — P. 295 – 306.

[Rat96]. Ratha N. K., Karu K., Chen S. and Jain A. A Real-Time Matching System for Large Fingerprint Databases//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1996. — Vol. 18. — N 8 (August). — P. 799 – 813.

[Rau92]. Rauber T., Steiger-Garcia A. 2-D Form Descriptions Based on a Normalized Parametric Polar Transformation (UNL-Transform). — Proc. MVA'92 — IARP Workshop on Machine Vision Applications. NEC Super Tower, Tokyo, Japan, December, 1992.

[Row98]. Rowley H., Baluja S. and Kanade T. Neural network-based face detection//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1998. — Vol. 20. — N 1 (January). — P. 23 – 37.



[San00]. Sanchez-Reillo R., Sanchez-Avila C., Gonzales-Marcos A. Biometric Identification through Hand Geometry Measurements//IEEE Trans. Pattern Analysis and Machine Intelligence. — 2000. — Vol. 22. — N 10 (October). — P. 1168 – 1171.

[Sch99]. Schreiner K. Biometrics: Prospects for going the distance.//IEEE Intelligent Systems. — 1999. — November/December. — P. 2 – 6.

[Sim00]. Sim T., Sukthankar R., Mullin M. S. Baluja. High-Performance Memory-based Face Recognition for Visitor Identification. Proceedings of Face and Gesture (forthcoming).

[Smi99]. Smiatacz M., Malina W. System Automatycznej Lokalizacji Obiektów Wykorzystujący Aktywne Modeli Kształtu. — Materiały XI Ogólnopolskiego Konwersatorium nt. Sztuczna Inteligencja — jej nowe oblicze. Siedlce, 1999, p. 179 – 186.

[Sol97]. Soldek J., Shmerko V., Phillips P., Kuchariew G., a.j. Image Analysis and Pattern Recognition in Biometric Technologies. — Proceedings of Fourth International Conference Pattern Recognition and Information Processing (PRIP'97), 20 – 22 May 1997, Minsk, Republic of Belarus, vol. 1, p. 362 – 378.

[Swe96]. Swets D. L., Weng J. Using Discriminant Eigenfeatures for Image Retrieval.//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1996. — Vol. 18. — N 8 (August). — P. 831 – 836.

[Swe99]. Swets D. L., Weng J. Hierarchical Discriminant Analysis for Image Retrieval.//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1999. — Vol. 21. — N 5 (May). — P. 386 – 400.

[Tha98]. Thalmann N., Karla R., Escher M. Face to virtual Face//Proceedings of The IEEE. — 1998. — Vol. 86. — N 5 (May). — P. 870 – 882.

[Tie97]. Tieng Q. M., Boles W. W. Wavelet-Based Affine Invariant Representation: A Tool for Recognizing Planar Objects in 3D Space//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 15. — N 8 (August). — P. 846 – 857.

[Tou74]. Tou J., Gonzalez R. Pattern Recognition Principles. Addison-Wesley Publishing Company, 1974. (Ту Дж., Гонсалес Р. Принципы распознавания образов. — М.: МИП. 1978).

[Tsa98]. Tsapatsoulis N., Alexopoulos V., Kollias S. A Vector Based Approximation of KLT and Its Application to Face Recognition, Proceedings of the IX European Signal Processing Conference EUSIPCO-98, September 8–11, 1998, Rodos Palace, Island of Rhodes, Greece, vol. III, p. 1581 – 1584.

[Tur91]. Turk M., Pentland A. Eigenfaces for Recognition//Journal of Cognitive Neuroscience. — 1991. — Vol. 3. — N 1 (January). — P. 71 – 86.

[Vet97]. Vetter T., Poggio T. Linear Objekt Classes and Image Synthesis from a Single Example Image.//IEEE Trans. Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — N 7 (Yuly). — P. 733 – 741.

[Wil97]. Wildes R. P. Iris Recognition: An Emerging Biometric Technology//Proceedings of The IEEE. — 1997. — Vol. 85. — N 9 (September). — P. 1348 – 1362.

[Woo97]. Woodward J. D. Biometrics: Privacy's Foe or Privacy's Friend//Proceedings of The IEEE. — 1997. — Vol. 85. — N 9 (September). — P. 1480 – 1491.

|             |         |
|-------------|---------|
| Предисл     |         |
| 1. Введ     | 1.1     |
|             | 1.2     |
| 2. Обзор    | 2.1     |
|             | 2.2.    |
|             | 2.3.    |
|             | 2.4. Т  |
|             | о       |
| 3. Элемент  | 3.1. Ф  |
|             | 3.2. С  |
|             | бс      |
|             | 3.3. Б  |
|             | 3.4. И  |
|             | 3.5. О  |
|             | ни      |
|             | 3.      |
|             | 3.      |
|             | 3.6. М  |
|             | 3.7. П  |
| 4. Предвари |         |
|             | 4.1. Вр |
|             | 4.1     |
|             | 4.1     |
|             | 4.1     |
|             | 4.2. Со |
|             | 4.3. Хо |
|             | 4.3     |
|             | 4.3     |



# О Г Л А В Л Е Н И Е

|   |            |
|---|------------|
| Предисловие .....   | 5          |
| <b>1. Введение в биометрику .....</b>   | <b>7</b>   |
| 1.1. Вводные понятия .....  | 8          |
| 1.2. Биометрические характеристики и их свойства .....                          | 13         |
| <b>2. Обзор некоторых биометрических технологий .....</b>                       | <b>18</b>  |
| 2.1. Технология отпечатков пальцев. История применения .....                    | 18         |
| 2.1.1. Информационные признаки отпечатков пальцев .....                         | 19         |
| 2.1.2. Стандарты на отпечатки пальцев .....                                     | 20         |
| 2.1.3. Устройства для получения отпечатков пальцев .....                        | 23         |
| 2.1.4. Сравнение отпечатков пальцев .....                                       | 27         |
| 2.1.5. Несколько слов о других подходах к реализации AFIS .....                 | 33         |
| 2.2. Технология, использующая геометрию руки .....                              | 38         |
| 2.3. Технология, использующая распознавание лиц .....                           | 42         |
| 2.3.1. Обзор методов распознавания лиц .....                                    | 44         |
| 2.3.2. База данных изображений лиц .....  | 46         |
| 2.3.3. Введение в методы PCA, ПКЛ и LDA .....                                   | 47         |
| 2.3.4. Отбор признаков при распознавании лиц .....                              | 52         |
| 2.3.5. Анализ локальных признаков .....   | 55         |
| 2.3.6. Использование эластичных моделей форм лица .....                         | 58         |
| 2.3.7. Коммерческие системы .....   | 61         |
| 2.3.8. Несколько слов о будущем .....   | 65         |
| 2.4. Технология, использующая распознавание по радужной оболочке глаза .....    | 67         |
| <b>3. Элементы цифровой обработки изображений .....</b>                         | <b>73</b>  |
| 3.1. Формы и способы представления цифровых изображений .....                   | 73         |
| 3.2. Система обозначений, используемая в алгоритмах обработки изображений ..... | 77         |
| 3.3. Базовые процедуры обработки изображений .....                              | 78         |
| 3.4. Инварианты геометрических преобразований .....                             | 85         |
| 3.5. Обработка изображений с использованием преобразований Фурье .....          | 93         |
| 3.5.1. Изображение и его преобразование Фурье .....                             | 93         |
| 3.5.2. Применение дискретного преобразования Фурье .....                        | 95         |
| 3.6. Метод главных компонент .....  | 103        |
| 3.7. Преобразование Фурье—Меллина .....   | 107        |
| <b>4. Предварительная обработка исходных данных .....</b>                       | <b>116</b> |
| 4.1. Выделение изображения лица, его деталей и параметров .....                 | 116        |
| 4.1.1. Выделение лица и его частей .....  | 116        |
| 4.1.2. Определение расстояния между центрами глаз .....                         | 122        |
| 4.1.3. Выделение изображения лица градиентными методами .....                   | 124        |
| 4.2. Согласование размеров изображений .....                                    | 125        |
| 4.3. Нормализация образов .....   | 130        |
| 4.3.1. Элиминирование поворотов .....   | 130        |
| 4.3.2. Элиминирование наклонов .....  | 132        |
|   | 239        |



|  |            |
|--|------------|
| 4.3.3. Нормализация яркости изображений . . . . .                                    | 135        |
| <b>5. Синтез системы распознавания для структурированных данных</b>                  | <b>141</b> |
| 5.1. Автоматическое разделение данных на обучающую<br>и контрольную группы . . . . . | 141        |
| 5.2. Отбор первичных признаков и их редукция . . . . .                               | 143        |
| 5.3. Уменьшение пространства признаков . . . . .                                     | 146        |
| 5.4. Процесс распознавания контрольных образов . . . . .                             | 149        |
| 5.5. Выбор параметров системы распознавания . . . . .                                | 151        |
| 5.5.1. Изменение числа классов и числа образов в классе                              | 151        |
| 5.5.2. Распознавание при наложении шума . . . . .                                    | 158        |
| 5.5.3. Распознавание при масштабировании образов . . .                               | 161        |
| 5.5.4. Распознавание при масштабировании и наложении<br>шума . . . . .               | 164        |
| 5.5.5. Распознавание образов, не принадлежащих собствен-<br>ной базе . . . . .       | 169        |
| <b>6. Системы распознавания «on-line»</b> . . . . .                                  | <b>172</b> |
| 6.1. Структура системы распознавания . . . . .                                       | 172        |
| 6.2. Виртуальные базы данных и выбор признаков . . . . .                             | 179        |
| 6.3. Синтез системы «Visitor Identification» . . . . .                               | 194        |
| 6.4. Системы, обеспечивающие контроль доступа . . . . .                              | 208        |
| 6.5. Интегрированные системы распознавания . . . . .                                 | 222        |
| Тестовая база данных (ORL Database of Faces) . . . . .                               | 225        |
| Словарь терминов * . . . . .   | 229        |
| Литература . . . . .   | 234        |



|       |     |
|-------|-----|
| ..... | 145 |
| ..... | 147 |
| ..... | 148 |
| ..... | 149 |
| ..... | 151 |
| ..... | 151 |
| ..... | 158 |
| ..... | 161 |
| ..... | 164 |
| ..... | 169 |
| ..... | 172 |
| ..... | 172 |
| ..... | 179 |
| ..... | 194 |
| ..... | 208 |
| ..... | 222 |
| ..... | 225 |
| ..... | 229 |
| ..... | 234 |





**TECHNICAL UNIVERSITY  
OF SZCZECIN  
FACULTY OF COMPUTER  
SCIENCE  
AND INFORMATION SYSTEMS**



Żołnierska Street 49, 71-210  
Szczecin,

Poland, tel./fax: (+48-91) 449-55-40, e-mail: [wi@wi.ps.pl](mailto:wi@wi.ps.pl),  
<http://www.wi.ps.pl>

Dean of Computer Science and Information Systems Faculty  
**prof. Jerzy Soldek**

---

***Employment:***

- 19 professors, • 27 doctors of science, • 61 masters of science,
- 36 persons of technical, administration and services staff
- The Faculty confers the degree of PhD in Computer Science

**Research teams:**

- |   |   |
|---|---|
| • Computer Graphics and Multimedia        | • Supercomputing and Parallel Programming   |
| • Computer Aided Decisions                | • Computer and Telecommunication Networks   |
| • Banking Information Technology          | • Cryptography and Software Engineering     |
| • Artificial Intelligence, Expert Systems | • Data Bases and Information Systems Design |
| • Image Recognition and Processing        | • Information Systems                       |
| • Multivalued Logic                       |   |
| • Synthesis of Reprogrammable Systems     |   |
| • Mathematical Modelling                  |   |

**Conferences organized by faculty:**

- Advanced of Computer Systems – ACS
- Pattern Recognition and Information Processing — PRIP
- Reconfigurability Unification & Customization — RUC





**TECHNICAL UNIVERSITY  
OF SZCZECIN  
FACULTY OF COMPUTER  
SCIENCE AND INFORMATION  
SYSTEMS**



Żołnierska Street 49,  
71-210 Szczecin,

Poland, tel./fax: (+48-91) 449-55-40, e-mail: [wi@wi.ps.pl](mailto:wi@wi.ps.pl),  
<http://www.wi.ps.pl>

Dean of Computer Science and Information Systems Faculty  
**prof. Jerzy Soldek**

---

**Diploma Courses (Pedagogic Specializations):**

- Software Engineering
- Microprocessor Systems
- Computer and Telecommunication Networks
- Computer Graphics and Multimedia
- Teledetection and Geographical Information Systems
- Information Systems
- Real Time Systems
- Systems of Computer Aided Decision
- Computer Systems in Administration
- Print and Publishing Systems

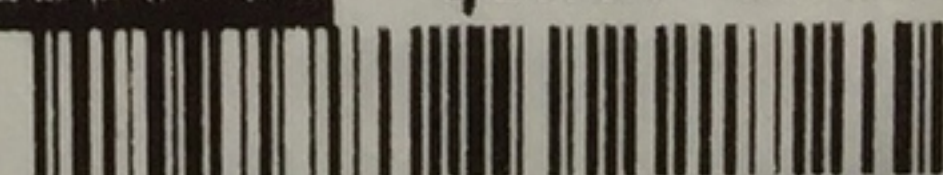
**Cooperation:**

- **France:** University of Technology (Troyes), Ecole Centrale de Lyon (Lyon) and Project STAR – Specialized Training in Aeronautic Research (Merignac, Bordeaux)
- **Canada:** Saint Francis Xavier University
- **Germany:** University of Goetting
- **Russian:** Electrotechnical University of St. Petersburg
- **Belarus:** Academy of Science and State University of Informatics and Radioelectronics

ГРУППА ОБЩЕСТВЕННЫЙ ДОМ КНИГ

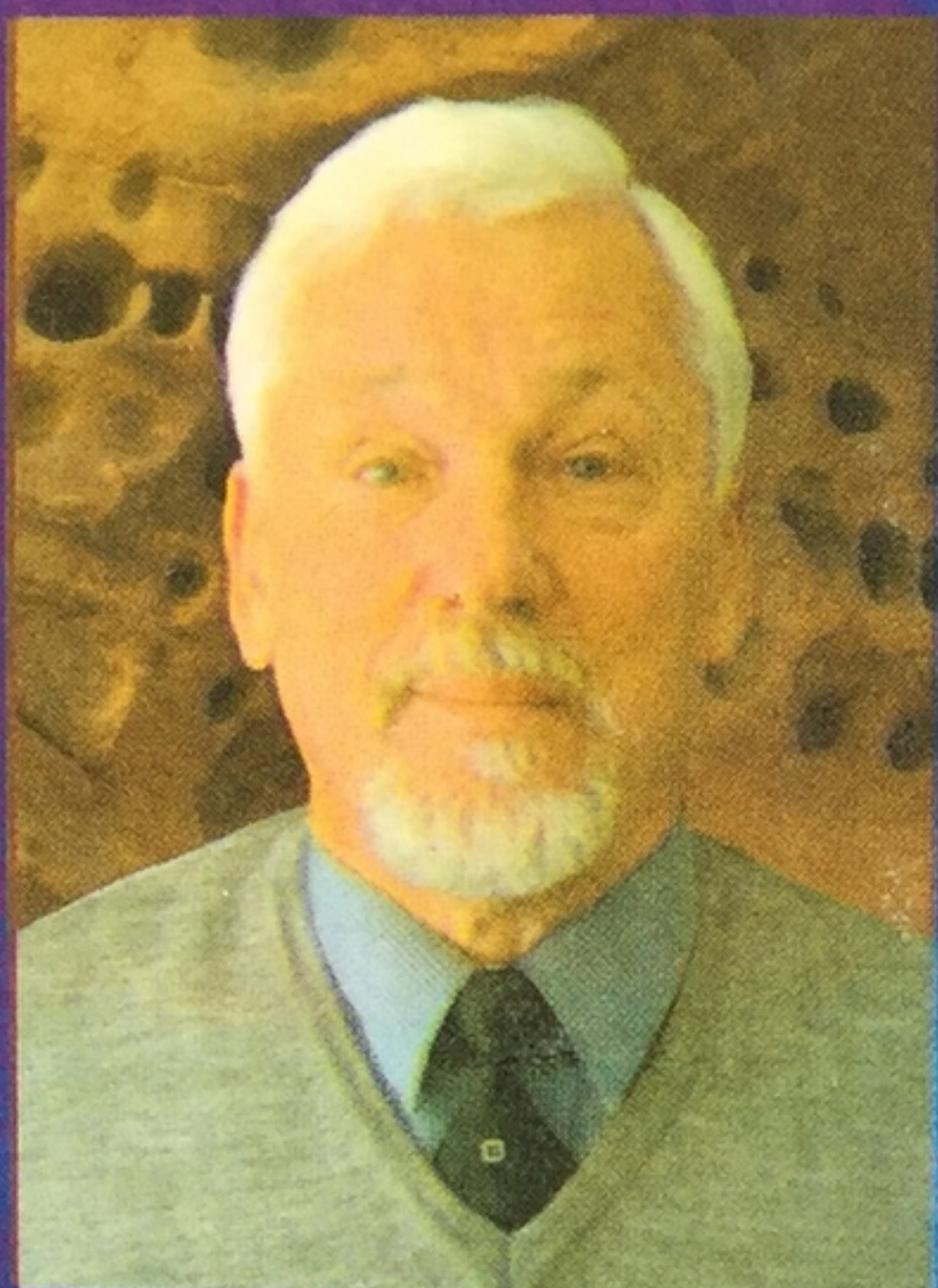
Кухарев Биометрические  
и системы

3253733 Цена: 170.00



20457106325373300010





**ГЕОРГИЙ АЛЕКСАНДРОВИЧ КУХАРЕВ** — доктор технических наук, с 1993 года профессор кафедры "Графики, обработки изображений и распознавания образов" на факультете информатики Щецинского политехнического университета (Польша).

Свою трудовую деятельность начал в 1960 году в Военно-механическом институте (Ленинград), где под руководством профессора С. А. Майорова отрабатывалась технология изготовления электронных

Длительное время сотрудничал с разработчиками ЭВМ серии "МИР" (Институт кибернетики АН УССР, Киев), являлся председателем технического Совета Ленинградского отделения Ассоциации пользователей ЭВМ серии "МИР" и членом Всесоюзного Совета этой Ассоциации.

В 1978 году успешно окончил аспирантуру Ленинградского института точной механики и оптики. Диссертацию на соискание ученой степени доктора технических наук защитил в 1986 году в Институте автоматики и вычислительной техники (Рига).

С 1988 по 1993 год работал в НПО "Океанприбор" (Санкт-Петербург), занимаясь разработкой бортовых вычислительных систем. Предложенные им технические решения защищены более чем 40 авторскими свидетельствами СССР на изобретения.

Г. А. Кухарев является автором (соавтором) нескольких монографий, посвященных методам и техническим средствам цифровой обработки сигналов и изображений. Среди них: "Адаптивное управление точностью обработки", "Быстрые дискретные ортогональные преобразования", "Систолические процессоры для обработки сигналов", "Алгоритмы и систолические процессоры для обработки многозначных данных", "Техника параллельной обработки бинарных данных на СБИС".

Настоящая монография "Биометрические системы" подготовлена на факультете информатики Щецинского политехнического университета в ходе международного сотрудничества по проблеме использования биометрических технологий в распознавании образов.

С 2000 года Г. А. Кухарев сотрудничает с лабораторией ICTT в Ecole Centrale de Lyon (Франция). Исследования последних лет посвящены проблеме детекции и локализации лиц людей на цветных изображениях с использованием представлений "skin-color".







Г. А. Рысбаев

# БІЛІМДІК ЕТІМДІК